



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 106

TO: Mark Navarro
Location: cm1/8A15
Art Unit: 1645
Monday, August 18, 2003

Case Serial Number: 09/904603

From: Barb O'Bryen
Location: Biotech-Chem Library
CM1-6A05
Phone: 308-4291

barbara.obryen@uspto.gov

Search Notes

O'Bryen, Barbara

From: Navarro, Albert
Sent: Wednesday, August 06, 2003 4:23 PM
To: O'Bryen, Barbara
Subject: 09/904,603

Mark Navarro
1645
306-3225
8A15

Hello Barb,

Could I trouble you for a search and an interference search of:

SEQ ID NO: 1-2

Thanks

Mark

2/24/97

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library 9th - Circ Desk



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2003, 06:52:34 ; Search time 2454 Seconds
(without alignments)
10669.177 Million cell updates/sec

Title: US-09-904-603-2

Perfect score: 640
Sequence: 1 CTCGCCGACGCCGACGCCGCC.....GCCGCCCTAGTCAGAGGCCCA 640

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBankl: *
1: gb_ba: *
2: gb_btq: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_com: *
21: em_ot: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pla: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	640	100.0	640	6	AR075096
2	640	100.0	640	6	AR155493
3	627.4	98.0	1030	9	AF276658
4	590.4	92.2	932	9	BC015810
5	552.4	86.3	912	9	HSM805143
6	435.2	68.0	984	10	BC010596
7	364	56.9	366	9	BT007452
8	297.8	46.5	1029	6	AX740487
9	296.2	46.3	15795	9	HSA344617
10	240.6	37.6	984	10	AF255953
11	240.2	37.5	66877	2	AC116062
12	239	37.3	190843	2	AC103943
13	238.5	37.3	861	6	AX402038
14	238.6	37.3	861	6	AX700303
15	238.6	37.3	861	10	RN005784
16	235.8	36.8	1632	5	BC043946
17	234.2	36.6	233682	2	AC096281
18	234.2	36.6	32693	2	AC132960
19	233	36.1	267290	2	AC134441
20	226.4	35.4	208868	2	AC124663
21	217.6	34.0	230614	10	AF626772
22	215.4	33.7	1910	6	AX740432
23	215.4	33.7	185479	2	AC073601
24	215	33.6	791	6	BD019774
25	215	33.6	791	6	BD099712
26	215	33.6	1937	9	AF183417
27	215	33.6	2110	9	AF303888
28	215	33.6	2151	9	BC018634
29	215	33.6	2155	9	AK025556
30	215	33.6	2157	9	AF087871
31	214.4	33.5	2159	6	AR650446
32	213.4	33.3	151111	9	AC125603
33	213.4	33.3	172393	2	AC025740
34	213.4	33.3	210933	2	AC027344
35	213.2	33.3	1805	5	BC049313
36	210.6	32.9	2073	9	HSM805336
37	201	31.4	465	6	BD111090
38	193	30.2	164047	10	AF69413
39	192.4	30.1	537	6	AX401525
40	191	29.8	2800	14	PT080885
41	178.8	27.9	269375	2	AC114657
42	174.2	27.2	318	5	CHRCDNA1AR
43	174.2	27.2	1181	9	AF276659
44	171	26.7	204150	2	AC140656
45	169.6	26.5	242177	2	AC120179

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
AR075096	Sequence 2 from patent US 5955312.	AR075096	AR075096.1	GI:10001848	Unknown.	Unknown.	Unclassified.	1 (bases 1 to 640)	Hillman, J.L. and Goli, S.K.	DNA encoding a novel microtubule-associated protein	Patent: US 5955312-A 2 21-SEP-1999;	Location/Qualifiers

source	1. .640	
	/organism="unknown"	
BASE COUNT	121 a 225 c 199 g	95 t
ORIGIN		

Query Match	100.0%;	Score 640;	DB 6;	Length 640;
Best Local Similarity	100.0%;	Pred. No. 1e-97;		
Matches 640;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 1 CTCCCGCAGCGCAGCGCGCGCGTGTCTGATCAGCGCGAGCGCCCGGAGCGCTTGACCGCGAGCGC 60

Db 1 CTCCCGCAGCGCGAGCGCGCGCGTGTCTGAGCGCGAGCGCCCGGAGCGCTTGACCGCGAGCGC 60

61 GGACCCCCGGAGCCCCCAACCGCAGACACATCCCCGGCCCCCAGAGCCCCGGCCCTGGCC 120
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573

Db

121 GCCCAGCGCCGGCGCGCATGCCCTTCACAGCAGCGGCGGAGCTTGC 180

181 CGACCGCCTTAAGGAGGTACAGCAGATCCCGGACCGAGCACCGCCAGCAAAATCCCGGTGAT 240

Db

241 CACGAGCCCTACAAGGTGAGAAGCAGCTGCCCTCTCGACAGACCAGTTTTCGT 300

Ov

301 CCCGGACCATCTCAACATATGACGCAGATTGGCAACAATATATCCCCCCCCCCCCTGA 360

DB	301	CCGGGACCATGTCAACATGAGCGGAGTTGGTCAAGTTCATTCGGGCCCGCGCTGCAGCTGAA	360
QY	361	CCCCACGCGAGGCTTCTCTCTGCTGCTGAGTGAACACAGACAGCATGGTGGTGTGTCCACGCC	420

Dbb 361 CCCCACGACAGGCTCTTCTCTCTGCTGCTGAACACAGACAGATGGTGTGTCCAGGCC 420

Oy 421 CATCGCGGACATCTACGAGCAGGAGAAAGACGAGAGACGGCTCTCTATATGTGTCTACGC 480

DB 421 CATTGGGGGACATCTACGACGAGGAAACAGAGACGGCTTCCCTCTATATGTCCTACGC 480

QY 481 CTCCGAGAAACCTTTGGGCTTCTGAGCCACGATGATGGGGGGCTCGGGCTGGGAGTCGGGG 540

Db 431 CTCCTCAGGAAACCTTGGGCTTCTGAGCCAGCACTAGGGGGGGCTGGGCTTGGAGTCGGGG 540

2Y 541 GGCCCGGGGACGGGCTGCCACAGAGACTTCTGGTTCCTGACATGAGTCGCTTACCGT 600

DB 541 GGGCCCCGGCAGGGCCCTGGCCAGAGAGCTTCCTGCTCCAGACTGAACAGCTAGCTGCCCCCTACCGT 600
 601 GGTGGGCTGGGCGAGCATGTGCCCCCTGTCAGAGGGCA 640
 QY

```

DU      001GGTGGGCTGGGCGAGGCAATGTCGCCCCCTAGTCAGAGGCA 640
RESULT 2

```

LOCUS	640 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 2 from patent US 6280733.			
ACCESSION	ARI65393			

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unknown.

REFERENCE
AUTHORS
TITLE
JOURNAL
VOLUME
PAGE
YEAR

1 (bases 1 to 640)
Hillman, J. L. and Golif, S. K.
Microtubule-associated protein ...
...
...
...
...

BASE COUNT	FEATURES	source
131	Location/Qualifiers	1..640
2235	/organism="unknown"	100..9
95		+

ORIGIN

Query Match	100.0%	Score 640;	DB 6;	Length 640;
Best Local Similarity	100.0%	Pred. No. 1e-97;		
Matches 640; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	CTCCCGAGCGGAGCGCGCGTGGTCAGGCGAGCGCCCGGAGGCCCTTGACGGCGAGGCGC	60
Db	1	CTCCCGAGCGGAGCGCGCGTGGTCAGCGGAGCGCCCGGAGGCCCTTGACGGCGAGGCGC	60

Qy 5 GGAGCCCCGAGCCCCCAACCGGAGACACATCCCCGGGGCCCCGAGCCCCGGGCTGGGC 120

61 GGAGCCCCGAGCCCCCAACCGGAGACACATCCCCGGGGCCCCGAGCCCCGGGCTGGGC 120

Db

Oy 121 GCCCAGCCGCGCCGCGCGATGCCCCCAGACCGCGGCTTTCAGCAGCAGCGCGAGCTTCGC 180
 Db 121 GCCCAGCCGCGCGCCGCGCGATGCCCCCAGACCGCGGCTTTCAGCAGCAGCGCGAGCTTCGC 180

QY
CGACCGCTGTAAAGGAGGTACAGCAATCCCGGACAGCACCACCAAAAATCCGGTGAT 240

DB
CGACCGCTGTAAAGGAGGTACAGCAGATCCCGACAGCACCAACCCCAATAATCCCGGTGAT 240

241 CATGAGCGCTACAAAGGCTGAGAGACAGCTGCGCCGCTCGAGACAAGACCAGTTTGGT 300

Db

301 CCGGACCATGTCAACATGAGCGAGTTGGTCAAGTCATCCGGCGCGCTTGACTGA 360

361 CCGGACCATGTCAACATGAGCGAGTTGGTCAAGTCATCCGGCGCGCTTGACTGA 420

421 CATCCGGGACATCTACGAGCAGGAGAAAGACGAGGACGGCTTCCCTATATGGGTACGC 480
 Db 361 CCCACGCGAGGCGCTTCTCTCGTGGTGAACACACACAGCATGGTGAAGTGTCCACGCC 420
 Ov 421

Db 421 CATTGGGGCATCTACGAGCAGGAGAAACGAGAGACGGCTCTCTAATATGGCTACGC 480

QY 481 CTCACAGAAACCTCGGCTTCTTAGCCAGCAGTAGGGGGGCTGGGCTGGGAGTCGGGG 540

Db 481 CTCCACGGAACCTTCGCTTCTGTAGCCACAGTAGGGGGGCTCGGCTTGGGAGTCGGGG 540

Qy 541 GGCCCCGGTCAGGGCCCTGCCACAGAGCTTCGTGTTCTCGAATCGAGCTGCCCTCACCGT 600

Db 541 GGCCCCGGTCAAGGCCCCCTGCCACAGAGAGCTTCTGGTTCTCGTAACTCAGCTGCCCTACCGT 600

Qy 601 GGTGGGCTGGGACGACATGTGCCCCCTTAGTCAGAGGCA 640

Db 601 GGTGGGCTGGGCAAGGCATGTGCCCCCCCTAGTCAGAGGGCA 640.

AF276658	1030 bp	mRNA	linear	PR1 16-APR-2003
LOCUS				
DEFINITION	Homo sapiens MAP1	light chain 3-like protein 1	mRNA, complete cds.	
ACCESSION	AF276658			

VERSION	1.1	61145023770
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
1 (pages 1 to 1030)
Yu, L.
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

JOURNAL. Submitted (09-JUN-2000) Institute of Genetics, Fudan University,
No. 220 Handan Road, Shanghai 200433, P.R. China
FEATURES Location/Qualifiers

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"
```

/note="LC3-like protein 1"

```

/codon_start=1
/product="MAP1 light chain 3-like protein 1"
/protein_id="AAK35151.1"
/db_xref="GI:13625771"
/translation="MPSDRPFKORRSFADRKKEVOQIDQHPKIPVIEEYKGEKUL
PVLQKTFELVDPDHVNMSELVKIRRRLOLNPTQAFLLVNOHSMVSVSPIDIVYEQ
KDEGFLYMYVAISOETFGF"
polya_signal
941..946
BASE COUNT      223 a      339 c      291 g      177 t
ORIGIN
Query Match      98.0%; Score 627.4; DB 9; Length 1030;
Best Local Similarity 99.7%; Pred. No. 1.2e-95;
Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 1 CTCGCCGACGCCGACGCCGCTGCTGACGCGGAGCGCCCGGAGCCCTTAGCGCGAGGCGC 60
    |||||||
DB 15 CTCGCCGACGCCGACGCCGCTGCTGACGCGGAGCGCCCGGAGCCCTTAGCGCGAGGCGC 74
    |||||||
OY 61 GGAG-CCCCGGAGCCCCCAAAACCGGAGACATCCCGCGCCGCGAGAGCCCGGCTGCG 119
    |||||
DB 75 GGAGCCCCGGAGCCCCCAAAACCGGAGACATCCCGCGCCCGGAGAGCCCGGCTGCG 134
    |||||
OY 120 CGCCGAGCGCGGCGCGCGATGCCCTCAGACCGGCGCTTTGAAGCAGCGGCGAGCTTGG 179
    |||||
DB 135 CGCCGAGCGCGGCGCGCGATGCCCTCAGACCGGCGCTTTGAAGCAGCGGCGAGCTTGG 194
    |||||
OY 180 CCGACCGCTGTAAAGAGGTACAGCAGATCCCGGAGCAGACCCCGGCAAAATCCCGGTGA 239
    |||||||
DB 195 CCGACCGCTGTAAAGAGGTACAGCAGATCCCGGAGCAGACCCCGGCAAAATCCCGGTGA 254
    |||||||
OY 240 TCATGACGCGCTACAAAGGTGAGAAAGCAGCTGCCGCTCTGGATCAGACCAAGTTTGG 299
    |||||||
DB 255 TCATGACGCGCTACAAAGGTGAGAAAGCAGCTGCCGCTCTGGATCAGACCAAGTTTGG 314
    |||||||
OY 300 TCCCGAGCAGATGTCAACATGAGCGAGTTGTCAAGATCATCCGCGCGCGCTGCAGCTGA 359
    |||||||
DB 315 TCCCGAGCAGATGTCAACATGAGCGAGTTGTGTCAAGATCATCCGCGCGCGCTGCAGCTGA 374
    |||||||
OY 360 ACCCGACCGAGCGCTTCTCTGCTGCTGGTGAACACAGACATGTTGAGTGTGTCCAGCG 419
    |||||||
DB 375 ACCCGACCGAGCGCTTCTCTGCTGCTGGTGAACACAGACATGTTGAGTGTGTCCAGCG 434
    |||||||
OY 420 CCATGCGCGAGCATCTACAGAGCAGAGAGAAAGACAGAGCGGCTTCTCTATATATGCTTACG 479
    |||||||
DB 435 CCATGCGCGAGCATCTACAGAGCAGAGAGAAAGACAGAGCGGCTTCTCTATATATGCTTACG 494
    |||||||
OY 480 CCTCCGAGAAACCTTCGGCTTCTGAGCCAGCAGTAGGGGGGGCTCGGCTTGGAGCTCGGG 539
    |||||||
DB 495 CCTCCGAGAAACCTTCGGCTTCTGAGCCAGCAGTAGGGGGGGCTCGGCTTGGAGCTCGGG 554
    |||||||
OY 540 GGGCGCGCGCTGAGCGCCCTGCCAGAGAGCTTCTGTTCTGAACTGAGCTGCTTACCG 599
    |||||||
DB 555 GGGCGCGCGCTGAGCGCCCTGCCAGAGAGCTTCTGTTCTGAACTGAGCTGCTTACCG 614
    |||||||
OY 600 TGGTGGCGTGGCAGAGCATGTGCCCCCTAGTCAGAGGGCA 640
    |||||||
DB 615 TGGTGGCGTGGCAGAGCATGTGCCCCCTAGTCAGAGGGCA 655
    |||||||

```

```

RESULT 4
LOCUS      BC015810                      932 bp      mRNA      linear      PRI 11-OCT-2001
DEFINITION Homo sapiens, MAP1 light chain 3-like protein 1, clone MGC:9521
IMAGE:3909192, mRNA, complete cds.
ACCESSION  BC015810
VERSION     BC015810.1  GI:16041838
KEYWORDS   MGC.
SOURCE      Homo sapiens (human)
ORGANISM   Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```

REFERENCE 1 (bases 1 to 932)
 Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (09-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-tr@mail.nih.gov
 Tissue Procurement: Afrc
 CDNA library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LUN1)
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Jhu, Chris Fjell, Erin Garland, Ken Glin,
 Leticia Hsiao, Martin Krzywinski, Keta Kutsche, Oliver Lee, Sue
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seefi, Jacqueline
 Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Zarrin.

FEATURES
 source
 1..932
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="LocusID:84557"
 /db_xref="taxon:9606"
 /clone="MGC:9521 IMAGE:3909192"
 /tissue_type="Olerus, leiomyosarcoma"
 /clone_lib="NTH_MGC_71"
 /lab_host="DH10B"
 /note="vector: pcMV-SpR16"
 104..469
 /codon_start=1
 /product="MAP1 light chain 3-like protein 1"
 /protein_id="AAH15810.1"
 /db_xref="GI:16041838"
 /translation="MPSDRPFKORRSFADRKKEVOQIDQHPKIPVIEEYKGEKUL
 PVLQKTFELVDPDHVNMSELVKIRRRLOLNPTQAFLLVNOHSMVSVSPIDIVYEQ
 KDEGFLYMYVAISOETFGF"
 KUEDGFLYMYVAISOETFGF

BASE COUNT 178 a 315 c 275 g 163 t
 ORIGIN

```

Query Match      92.2%; Score 590.4; DB 9; Length 932;
Best Local Similarity 99.7%; Pred. No. 1.9e-89;
Matches 602; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 38 CGAGCCCTTGAGCGCGGAGCGCGGAG-CCGGGAGCGCGGCAACCGGAGATATATATAT 95
    |||||||
DB 1 CGAGCCCTTGAGCGCGGAGCGCGGAG-CCGGGAGCGCGGCAACCGGAGATATATATAT 60
    |||||||
OY 97 GCGCCGAGAGCCCGGAGTGTGGGCGGAGCGGAGCGGAGCGGAGTGTGGGCGGAGTGTGG 156
    |||||||
DB 61 GCGCCGAGAGCCCGGAGTGTGGGCGGAGCGGAGCGGAGCGGAGTGTGGGCGGAGTGTGG 120
    |||||||
OY 157 TTTCAGCAGCGCGCGGAGTGTGGGAGCGGCTGTGAAGCAGGTACAGCAGATATATATAT 216
    |||||||
DB 121 TTTCAGCAGCGCGCGGAGTGTGGGAGCGGCTGTGAAGCAGGTACAGCAGATATATATAT 189
    |||||||
OY 217 GCAGCCGAGCAAAATCCCGGATGTAGAGCGTACAAAGGAGGAGAGCAGTATATATATAT 276
    |||||||
DB 181 GCAGCCGAGCAAAATCCCGGATGTAGAGCGTATAGAGCGTATAGAGGAGTATATATATAT 240
    |||||||
OY 277 CTTGACCAAGGCAAGTTTTCGATCCGAGACCATGTCAACATGAGCGAGTGTATATAT 335
    |||||||

```

Db	241	CCTGACACAAGACCAAGTTTGTGTGTCCTCCGGAGCACCATGTCACACATAGCGAGTGTGTCAAAGAT	300
OY	337	CATCGGGGGCCGCCGAGCGTGAACCCACACGGCAGAGCTTCTTCTCTCTGTGTAAACCAGCA	396
Db	301	CATCCGGGGCCGCCCTGACAGCTGAACCCACACGAGGCTTCTTCTCTCTGTGTAAACCAGCA	360
OY	397	CAGCATGGTGAAGTGTGTCTCCACGCCCATCGCGGACATCTTCAGAGCAGAGAAAGACGAGGA	456
Db	361	CAGCATGGTGAAGTGTGTCTCCACGCCCATCGCGGACATCTTCAGAGCAGAGAAAGACGAGGA	420
OY	457	CGGCTTCTCTATATATGTCCTACGCCCTCCAGGAAACCTTCGGGCTTCTTGAGCCAGCAGTAG	516
Db	421	CGGCTTCTCTATATATGTCCTACGCCCTCCAGGAAACCTTCGGGCTTCTTGAGCCAGCAGTAG	480
OY	517	GGGGGCTCGGGCTGGGAGTCGGGGGGGGCCCGGCTCGAGGCCCTGCCAGAAAGCTTCTTGATT	576
Db	481	GGGGGCTCGGGCTGGGAGTCGGGGGGGGCCCGGCTCGAGGCCCTGCCAGAAAGCTTCTTGATT	540
OY	577	CCTGAAGCTGAGCTGGCTCTACCGTGGTGGGCTGGGACAGGCAATGTGCCCCCTTAGTCAGAG	636
Db	541	CCTGAAGCTGAGCTGGCTCTACCGTGGTGGGCTGGGACAGGCAATGTGCCCCCTTAGTCAGAG	600
OY	637	GGCA 640	
Db	601	GGCA 604	

LOCUS	DEFINITION	Accession
RESULT 5 HSM805143		
HSM805143	912 bp mRNA linear	PRI 12-JUL-2002
	Homo sapiens mRNA; cDNA DKFZp761I0515 (from clone DKFZp761I0515); complete cds.	

VERSION	AL8333855.1	GI:21739336
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (phases 1 to 912)	Ansoerge,W., Wirkner,U., Mewes,H.W., Weill,B. and Wiemann,S.	Direct Submission	Submitted (10-JUL-2002)	GSF, Institut fuer Bioinformatik MIPS, Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany
				Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

German Genome Project. This clone (DKRzp/61L0515) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcententrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cGNA>.

FEATURES	Location/Qualifiers
source	1. .912

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:.9606"
/chromosome="20"
/clone="DKFZp761L0515"
/tissue_type="amygdala"
/clone_id="761 (synonym:
DH10B; sites NotI + SalI"
/dev_strage="adult"
1..912
./gene="DKFZp761L0515"
66..431
gene
CDS

```

```

polyA_signal
polyA_site
BASE COUNT      193 a      298 c      260 g      161 t
ORIGIN

```

Query Match	Best Local Similarity	Score	DB	Length
Matches	564: Conservative	96.3%: 89.6%: 99.6%:	0: Mismatches	1: Indels
QY	75	CCCAAACCGGACAGACATCCCCGGCGCCCGACAGCCCGGCGCTGCGCGCCCGACCGGGCCC	134	
Db	1	CCCCAAACCGGACAGACATCCCCGGCGCCCGACAGCCCGGCGCTGCGCGCCCGACCGGGCCC	60	
QY	135	GCGCGATGCCCTCGAGACCGGCGCTTTCAAGACAGCGCGGAGCTTCGCGACCGCTTAAAG	194	
Db	61	GCGCGATGCCCTCGAGACCGGCGCTTTCAAGACAGCGCGGAGCTTCGCGACCGCTTAAAG	120	
QY	195	AGGTACAGACATCCGGCAGCAGACCCGACGAATATCCGCTGATTCATCGAGCCCTACA	254	
Db	121	AGGTACAGACATCCGGCAGCAGACCCGACGAATATCCGCTGATTCATCGAGCCCTACA	180	
QY	255	AGGCTGAGACAGCTGCGCGCTCTCGACAGACCAATTTTGGTCCCGGACATGTCA	314	
Db	181	AGGCTGAGAACAGCTGCGCGCTCTCGACAGACCAATTTTGGTCCCGGACATGTCA	240	
QY	315	ACATGAGAGGATTTGGTCAAGATTCATCCGGCGCGGCTGACCTGTAAGCCGACGCGGCT	374	
Db	241	ACATGAGAGGATTTGGTCAAGATTCATCCGGCGCGGCTGACCTGTAAGCCGACGCGGCT	300	
QY	375	TCCTTCCTGCTGCTGTAACCCAGCAGACATGTGAGTGTGTCCACGCGCCATCGCGGACATCT	434	
Db	301	TCCTTCCTGCTGCTGTAACCCAGCAGACATGTGAGTGTGTCCACGCGCCATCGCGGACATCT	360	
QY	435	ACGAGCAGGAGAAAGACGAGCAGGCGTCTCTATATGCTGTACGCGCCCGCAGGAACTT	494	
Db	361	ACGAGCAGGAGAAAGACGAGCAGGCGTCTCTATATGCTGTACGCGCCCGCAGGAACTT	420	
QY	495	TCGGCTTCTGAGCCAGCAGTATGAGGGGGCTCGGCTGGAGTGGGGGGGCCCGGTCAGGC	554	
Db	421	TCGGCTTCTGAGCCAGCAGTATGAGGGGGCTCGGCTGGAGTGGGGGGGCCCGGTCAGGC	480	
QY	555	CCTCGCCACGAGAGCTTCTGCTTCCTGGAAGTACGCTCTAACCGTGGTGGGCTGGGAG	614	
Db	481	CCTCGCCACGAGAGCTCTGCTTCCTGGAAGTACGCTCTAACCGTGGTGGGCTGGGAG	540	
QY	615	GCATGTGCCCCCTAGTCAGAGGCA	640	
Db	541	GCATGTG-CCCCCTAGTCAGAGGCA	565	

RESULT 6	BC010596	984 bp	mrna	linear	ROD 16-APR-2003
LOCUS	BC010596				
DEFINITION	Mus musculus RIKEN CDNA 4922501H04 gene, mRNA (cdna clone MGC:6325				
ACCESSION	BC010596				
VERSION	BC010596.1	GI:14714881			
KEYWORDS	MGC.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				

ORGANISM	REFERENCE
<i>Mus musculus</i>	Klausner, R. D., Feingold, E. A., Grouse, L. H., Dege, J. G., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Mammalia: Eutheria; Rodentia; Scutirognathi; Muridae; Murinae; Mus.
<i>Eukaryota</i> ; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	1 (bases 1 to 984)

Hopkins, R.F., Jordan, H., Moore, T., Max, S.L., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Schaefer, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshyuk, I.S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Guaratine, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kellman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Boniford, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smolus, D.F., Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
MEDLINE
PUBMED
22188257
12477932
2 (bases 1 to 984)

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (10-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Guaratine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisgeq, H., Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Navevali, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 6 Row: 9 Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, similarity but not identity to protein.

FEATURES
SOURCE
location/qualifiers
1..984
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="MGC:6325 IMAGE:3256801"
/tissue_type="Mammary tumor. WAP-Tag model. 5 months old.
gross tissue."
/clone_id="DH10B"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..984
/gene="4922501H04Rik"
/db_xref="LocustID:66734"
/db_xref="MGI:1913984"
144..509
/codon_start=1
/product="RIKEN cDNA 4922501H04"
/protein_id="AAH10596.1"
/db_xref="GI:14714882"
/db_xref="LocustID:66734"

Query Match 68.0%; Score 415.2; DB 10; Length 984;
Best local Similarity 84.0%; Pred. No. 1.9e-63;
Matches 531; Conservative 0; Mismatches 88; Indels 13; Gaps 3;
BASE COUNT 208 a 320 c 271 g 185 t
ORIGIN
/translation="MPSURPEKRSFADCKEVOQIIRDOHPKIPVLEPKGEKUL
PVLUKTEVPHVHVMSELYKLTIRRIJUNTOAFLLVNOHSMVNSVSLPAIDYEDP
KDEGEFLMYVASSETFSP"

3 CCGGAGCGCGAGCGCGCGTCTAGCGCGAGCGCGCGAGCGCTTTGAGCGCGAGCGCGCG 62
1 CCGGAGCGCGCGAGCGCGCGAGCGCGAGCGCGCGAGCGCGCGCTTGAAGCGCGCGCGCG 60
63 AGCGCGCGAGCGCGCGCGAGCGCGAGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 116
61 AGCGCGCGCGCGCGCGCGAGCGCGAGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 120
117 GCGCGCGCGAGCGCGCGCGCGCGAGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 176
121 GCGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
177 TCGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 236
181 TCGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
237 TGATCATGAGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 296
241 TGATCATGAGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
297 TGGTCCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 356
301 TGGTCCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
357 TGAACCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 416
361 TGAACCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
417 GCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 476
421 CAGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
477 AGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
481 AGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
537 GAGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 595
535 TGGGAGGTTCGTGAGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCG 594
596 AGCGGTGTGCGGTGGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 627
595 TGGCGAGGTGATGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 626

RESULT 7
BT007452
LOCUS
DEFINITION
Homo sapiens microtubule-associated protein 1 light chain 4 alpha
mRNA, complete cds.
BT007452.1 GI:30584712
ACCESSION
BT007452
VERSION
BT007452.1
KEYWORDS
FT:CDNA
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Gnathata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 366)
Kalinine, N., Chen, X., Kollis, A., Haller, A., Hines, L., Eisenstein, S.,
Koundina, M., Kaprielian, J., Moreira, D., Kelley, T., Laber, J., Liu, Y.,
Pheasant, M. and Farmer, A.
Cloning of human full-length cDNA in BD Creator(TM) System Insert
vector

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 366)
AUTHORS Kaindl,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labber,J., Lin,Y., Phelan,M. and Farmer,A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the PDR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES
Source location/Qualifiers
1..366
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00946X1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: PDR-Dual"
1..366
/codon_start=1
/product="microtubule-associated protein 1 light chain 3 alpha"
/protein_id="AAP36120.1"
/db_xref="GI:30583743"
/translation="MPSDRPFKORSRFADRCKEYOIRDOHPSKIPYIEERYKGEKOLPVLDKTEFVDPDHVNMSEIVYKIIIRRLQNLNTQAFLLVNHSHSVSTPLADIYEDQKDEGFVLYMYASQETFGF"
BASE COUNT 84 a 116 c 101 g 65 t
ORIGIN

Query Match 56.9%; Score 364; DB 9; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 ATGCCCTCAGACGCGCTTTTCAGCAGCGGCGAGCTTGGCCGAGCGCTGAAGAGGTA 199
DB 1 ATGCCCTCAGACGCGCTTTTCAGCAGCGGCGAGCTTGGCCGAGCGCTGAAGAGGTA 60-
QY 200 CAGCAGATCCGCGACGACGACCCGCAAAATCCCGGATCATCGACGCTACAAAGGT 259
DB 61 CAGCAGATCCGCGACGACGACCCGCAAAATCCCGGATCATCGACGCTACAAAGGT 120
QY 260 GAGAAGCAGCTGCCCGCTCTGGACAGACCAAGTTTGGTCCCGGACCATGTCAACATG 319
DB 121 GAGAAGCAGCTGCCCGCTCTGGACAGACCAAGTTTGGTCCCGGACCATGTCAACATG 180
QY 320 AGCAGTTGGTCAAGATCATCCGCGCGCTGACGTGAACCCACGACGCGCTTCTTC 379
DB 181 AGCAGTTGGTCAAGATCATCCGCGCGCTGACGTGAACCCACGACGCGCTTCTTC 240
QY 380 CTGCTGGGAACGACGACGATGAGTGTCTCCACGCCCATTCGGGAGATATACAGAG 439
DB 241 CTGCTGGTGAACGACGACGATGAGTGTCTCCACGCCCATTCGGGAGATATACAGAG 300
QY 440 CAGGAGAAGACGAGAGCGCTTCTATATGTTAGCTACGCTCCAGGAACCTTGGC 499
DB 301 CAGGAGAAGACGAGAGCGCTTCTATATGTTAGCTACGCTCCAGGAACCTTGGC 360
QY 500 TTCT 503
DB 361 TTCT 364

RESULT 8
AX740487 1029 bp DNA Linear PAT 10-MAY-2003
LOCUS AX740487
DEFINITION Sequence 76 from Patent WO02020756.
ACCESSION AX740487
VERSION AX740487.1 GI:30523651
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Stuart,J., Lincoln,S.E., Altus,C.M., Dufour,G., Chalup,M., Hillman,J.L., Jones,A.L., Yu,J.Y., Wright,R.J., Gietzen,D., Liu,T., Yap,P., Dahl,C.R., Momiyama,M.G., Bradley,D., Rohatgi,S., Harris,B., Roseberry, Ann.M., Gestin,E.H., Peralta,C.H., David,M.H., Panzer, Scott.R., Flores,V., Daffo,A., Marwaha,R., Chen,A.J., Chang,S.C., Au,A.P. and Iman,R.R.
Secretory molecules
Patent: WO 02020756-A 76 14-MAR-2002;
Incyte Genomics, Inc. (US)

TITLE Location/Qualifiers
JOURNAL 1..1029
FEATURES
Source /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LG:247776.14:2000SEP08"
BASE COUNT 172 a 318 c 354 g 183 t 2 others
ORIGIN

Query Match 46.5%; Score 297.8; DB 6; Length 1029;
Best Local Similarity 99.3%; Pred. No. 2e-40;
Matches 299; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 340 CCGGCGCGCGCTGACGTCGACGACCCACGACGCGCTTCTCTCGTGTGACACCCACAG 399
DB 554 CAGGCGCGCGCTGACGTCGACGACCCACGACGCGCTTCTCTCGTGTGACACCCACAG 613
QY 400 CATGCTAGTGTGTCCACGCGCCATCGCGAGCATCTACGACGAGAGAGAGAGAGAG 459
DB 614 CATGCTAGTGTGTGTCCACGCGCCATCGCGAGCATCTACGACGAGAGAGAGAGAGAG 673
QY 460 CTTCCTCTATATGTGTCTACGCTTCCAGGAACCTTGGCTTGTGACGACGAGAGAG 519
DB 674 CTTCCTCTATATGTGTCTACGCTTCCAGGAACCTTGGCTTGTGACGACGAGAGAG 733
QY 520 GGCTCGGCGCTGGAGTGGGGGGGCGCGGTGAGGCGCCGACGAGAGCTTGTGTTCT 579
DB 734 GGCTCGGCGCTGGAGTGGGGGGGCGCGGTGAGGCGCCGACGAGAGCTTGTGTTCT 793
QY 580 GAACGTGACCTGCTCTACCTGCTGTGGCTGGGACAGCATGTGCCCTTACTAGAGAGGC 639
DB 794 GAACGTGACCTGCTCTACCTGCTGTGGCTGGGACAGCATGTGCCCTTACTAGAGAGGC 853
QY 640 A 640
DB 854 A 854

RESULT 9
HSA346K17 157975 bp DNA Linear PRI 19-OCT-2000
LOCUS HSA346K17
DEFINITION Human DNA sequence from clone Rpl1-346K17 on chromosome 20 Contains a novel gene encoding a protein similar to the cell division control protein 91 (CDC91) from yeast, a novel gene encoding two isoforms similar to MAP1A3 (microtubule-associated proteins 1A/1B lightchain 3) from Rat, 3 Cpg islands, ESTs, STSs and GSSs, complete sequence.
ACCESSION AL118520
VERSION AL118520.26 GI:9801544
HTG: CDC91; Cpg island; microtubule-associated proteins.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.
1 (bases 1 to 157975)

AUTHORS Sehtra,H.

TITLE Direct Submission

JOURNAL Submitted (19-OCT-2000) Sanger Centre, Hinxton, Cambridgeshire, UK10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
Requests: clonerquest@sanger.ac.uk

COMMENT On Aug 16, 2000 this sequence version replaced gi:59717213.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; SW: SWISSPROT; Tr: TREMBL; Wp: WormPEP; Information on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20

RP11-346K17 is from the library RPC1-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/VECTOR:pbac3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-346K17. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP11-346K17 is at 1 in this sequence. The true left end of clone RP5-1181N3 is at 157876 in this sequence. The true right end of clone RP5-914B9 is at 92396 in this sequence.

FEATURES

SOURCE

1. 157975
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-346K17"
/clone_1id="RPC1-11.2"
19. 630
/note="match: GSS: Em:A0541128"
86. 507
/note="match: GSS: Em:A0093367"
2783. 2959
/note="MLT17 repeat: matches 120. 292 of consensus"
2964. 16247
/gene="ba346K17.1"
join(2964..2991,5980..6104,15283..15338,15422..15528,15811..16247)
/gene="ba346K17.1"
/product="ba346K17.1.1 (Novel protein similar to MAP1ALC3 (microtubule-associated proteins 1A/1B light chain 3) from Rat, isoform 1)"
/note="match: ESTs: Em:HI8312 Em:AW480459 Em:W82731 Em:AW161970 Em:AA016389 Em:AA197682 Em:AA495746 Em:AW161293"
/evidence=not_experimental
3543. 3896
/note="THRC repeat: matches 1. 369 of consensus"
misc_feature
complement(5105..5567)
/note="match: GSS: Em:A0834732"
complement(5177..5581)
/note="match: GSS: Em:A0132454"

repeat_region
CDS

5554. 5943
/note="MER6A repeat: matches 27. 438 of consensus"
join(6053..6104,15283..15338,15422..15528,15811..15973)
/gene="ba346K17.1"
/note="match: proteins: Tr:094272 Tr:09SL04 Sw:J87066 Tr:09SLN6 Sw:090490 Sw:062625 Sw:041515"
/codon_start=1
/evidence=not_experimental

repeat_region

/product="ba346K17.1.1 (Novel protein similar to MAP1ALC3 (microtubule-associated proteins 1A/1B light chain 3) from Rat, isoform 1)"
/protein_id="CAC14078.1"
/db_xref="GI:10944272"
/translation="MKMKPFSSPGCAAVPADCKREVGQIRGHPKIVLEKYG EKQIPVLDKTKRIVPDVNNSELVKTIIRRLQINPTQAFPIILVQHSVSVIPIALV EDERDDEGLMYTASQELFGF"
639. 7044
/note="L1ME2 repeat: matches 5423. 5826 of consensus"
7334. 7542
/note="L1ME2 repeat: matches 5825. 6044 of consensus"

repeat_region

7835. 7946
/note="L1ME2 repeat: matches 6044. 6150 of consensus"
complement(9544..10055)
/note="match: GSS: Em:AV141567"
complement(9583..10030)

misc_feature

/note="match: GSS: Em:AV186762"
10052. 10443
/gene="ba346K17.1"
/note="match: GSS: Em:AV246853"
10102. 10183
/note="L1P3 repeat: matches 5503. 5586 of consensus"

misc_feature

11083. 11134
/note="26 copies 2 mer lt 73% conserved"
11729. 11890
/note="FRAM repeat: matches 1. 161 of consensus"

repeat_region

repeat_region

repeat_region

misc_feature

CDS

14407. 15589
/gene="ba346K17.1"
/note="CpG island"
/evidence=not_experimental
join(614797..14583,15283..15338,15422..15528,15811..15973)
/gene="ba346K17.1"
/note="match: proteins: Sw:J87068 Sw:062625 Tr:09SL04 Tr:094272 Sw:041515"
/codon_start=1
/evidence=not_experimental

/product="ba346K17.1.2 (Novel protein similar to MAP1ALC3 (microtubule-associated proteins 1A/1B light chain 3) from Rat, isoform 2)"
/protein_id="CAC14079.1"
/db_xref="GI:10944273"
/translation="ELTPVAAAVALASISQALEPAKPSKQIADISPRKAVCAV SKARAPSDRFKORSRADCKEVQVQIRGHPKIVITIRYGRKQLVLIKIKELVVDASQNNSELVKTIIRRLQINPTQAFPIILVQHSVSVIPIALV EDERDDEGLMYTASQELFGF"
15345. 15412
/note="34 copies 2 mer cc 66% conserved"

repeat_region

gene

complement(16618..133160)
/gene="ba346K17.2"
complement(join(16618..17054,31179..31321,37623..37747,41512..41655,44528..44682,72117..72214,96630..96730,93951..94060,100239..100301,101350..101409,113255..113319,133031..133160))
/gene="ba346K17.2"
/product="ba346K17.2 (A novel protein similar to the cell division control protein 91 (CPC91, YLR459W or L9122.2) from Yeast)"

/note="A novel protein similar to the cell division control protein 91 (CPC91, YLR459W or L9122.2) from Yeast"
/evidence=not_experimental

polyA_site

polyA_signal

complement(16618)
/gene="ba346K17.2"
complement(16638..16643)
/gene="ba346K17.2"

CDS
 complement(join(16941..17054,31179..31321,37623..37747,
 41512..41655,44528..44682,72117..72214,90690..90790,
 93951..94060,100239..100301,101350..101409,113255..113319,
 133031..133160))
 /gene="ba346k17.2"
 /note="match: ESTs: Em:AI315824 Em:AW753001 Em:AW046539
 Em:AA317421 Em:AA157236 Em:AN354502
 match: proteins: Sw:PA1733"
 /codon_start=1
 /evidence=not_experimental
 /product="ba346k17.2 (A novel protein similar to the cell
 division control protein 91 (CDC91, YLR459W or L9122.2)
 from Yeast)"
 /protein_id="CAC14080.1"
 /db_xref="GI:10944274"
 repeat_region
 /note="11MC3 repeat: matches 6991..7739 of consensus"
 21784..21835
 /note="26 copies 2 mer ac 88% conserved"
 complement(21837..22058)
 /gene="ba346k17.2"
 /note="match: GSS: Em:AQ061565"
 complement(21841..22045)
 /gene="ba346k17.2"
 /note="match: GSS: Em:AQ415451"
 21906..21949
 /note="22 copies 2 mer ca 75% conserved"
 repeat_region
 22079..22186
 /note="54 copies 2 mer ac 66% conserved"
 repeat_region
 23863..23908
 /note="23 copies 2 mer tc 82% conserved"
 repeat_region
 28237..28304
 /note="34 copies 2 mer tt 66% conserved"
 29295..29742
 misc_feature
 46.3%: Score 296.2; DB 9: Length 157975;
 Best Local Similarity 99.0%; Pred. No. 1.7e-40;
 Matches 298; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

340 CCGGCGCGCGCTGACGTGAACCCACGAGCGCTTCTTCCGTGCTGTAACACACAG 399
 |||||||
 15808 CAGGCGCGCGCTGACGTGAACCCACGAGCGCTTCTTCCGTGCTGTAACACACAG 15867
 |||||||
 400 CATGCTGAGTGTGTCCAGCGCATGCGGACATCTACAGACAGAGAAGAGGAGCG 459
 |||||||
 15868 CATGCTGAGTGTGTCCAGCGCATGCGGACATCTACAGACAGAGAAGAGGAGCG 15927
 |||||||
 460 CTTCCTATATATGCTCTACGCGCTCCAGGAACCTTGGCTTCTGAGCAGACAGT 519
 |||||||
 15928 CTTCCTATATATGCTCTACGCGCTCCAGGAACCTTGGCTTCTGAGCAGACAGT 15987
 |||||||
 520 GGTCTGCGCGCTGAGAGTGGGGGGCCCGGTGACAGCGCTCCAGAGACCTTCTG 579
 |||||||
 15988 GGTCTGCGCGCTGAGAGTGGGGGGCCCGGTGACAGCGCTCCAGAGACCTTCTG 16047
 |||||||
 580 GAATGAGCTGCGCTCTACCGTGTGGGTGGGAGGCAWGTGCCCCCTAGTACAGAGGC 639
 |||||||
 16048 GAATGAGCTGCGCTCTACCGTGTGGGTGGGAGGCAWGTGCCCCCTAGTACAGAGGC 16107
 |||||||
 640 A 640
 16108 A 16108

AF255953
 LOCUS AF255953 984 bp mRNA linear ROD 02-MAR-2002
 DEFINITION Mus musculus MAP1A/1B light chain 3 subunit mRNA, complete cds.
 ACCSSION AF255953
 VERSION AF255953.1 GI:19070136
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 984)
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-2000) Institute of Genetics, Fudan University,
 No. 220 Handan Road, Shanghai 200433, P.R. China
 FEATURES
 source
 1..984
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 76..453
 /codon_start=1
 /product="MAP1A/1B light chain 3 subunit"
 /protein_id="AAL83723.1"
 /db_xref="GI:19070137"
 /translation="MPSEKTFKRRSFQEDVDRLIREQHPKIPVIERYKGEKOL
 PVLDTKFLVDPHVNMSSELKILIRRLQLANQAFLLVNGHSWVSVPISVEYSE
 RDEDFLVVAVSOEPTGAMAV"
 BASE COUNT 260 a 260 c 256 g 208 t
 ORIGIN

Query Match
 Best Local Similarity 37.6%; Score 240.6; DB 10: Length 984;
 Matches 300; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

100 CCCAGAGCCCGCGCTGCGCCCGACCGCGGCGCGGATGCGCTTACAGCGCCCTT 159
 |||||
 36 CGCCGAGTCAAGATGCTGCTGCGGACCGCGCGCAAGCGCTCGAAGAACCTT 95
 |||||
 160 CAAGCAGCGCGAGCTTCCGCGCGCTGTAAGGAGTACAGATCCGCGACACAGA 219
 |||||
 96 CAAGCAGCGCGAGCTTCCGCGCGCTGTAAGGAGTACAGATCCGCGACACAGA 155
 |||||
 220 CCCAGCAAAATCCCGGTGATCAGAGCGCTTACAGAGGAGTACAGAGCTGCGTCT 279
 |||||
 156 CCCAGCAAAATCCCGGTGATCAGAGCGCTTACAGAGGAGTACAGAGCTGCGTCT 215
 |||||
 280 GGACAAAGCAAGTTTGTGTCGCGGACCATGTCAACATGAGCGAGTTGTCAGATCAT 339
 |||||
 216 GGACAAAGCAAGTTTGTGTCGCGGACCATGTCAACATGAGCGAGTTGTCAGATCAT 275
 |||||
 340 CCGGCGCGCGCTGACGTGAACCCACGAGCGCTTCTTCCGTGCTGTAACACACAG 399
 |||||
 276 CAGAGCGCGCTTCCGCGCGCTTCCGCGCGCTTCTTCCGTGCTGTAACACACAG 335
 |||||
 400 CATGCTGAGTGTGTCCAGCGCATGCGGACATCTACAGACAGAGAAGAGGAGCG 459
 |||||
 336 CATGCTGAGTGTGTCCAGCGCATGCGGACATCTACAGACAGAGAAGAGGAGCG 395
 |||||
 460 CTTCCTATATATGCTCTACGCGCTCCAGGAACCTTGGCTTCTGAGCAGACAGT 498
 |||||
 396 CTTCCTATATATGCTCTACGCGCTCCAGGAACCTTGGCTTCTGAGCAGACAGT 434
 |||||

RESULT 11
 LOCUS AC116062 66877 bp DNA linear HTG 13-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-76N21, WORKING DRAFT SEQUENCE, 38
 AC116062
 AC116062
 AC116062.6 GI:30578911
 VERSION
 KEYWORDS
 HTGS_PHASE1: HTGS_DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eumariota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 66877)

Muzny,D,Marle, Metzker,M, Lee, Abramzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Arduano,D, Ayalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,K, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Kocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,I, Garza,M, Geregorgis,E, Geer,K, Gill,R, Grady,M, Guerra,M, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogue,M, Hollins,B, Howells,S, Hulik,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Kapachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorensunwa,L, Louiseged,H, Lozado,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindartine,M, Mahmoud,M, Malloy,K, Mangun,A, Mangun,B, Mapua,P, Martin,K, Martin,R, Martinez,E, McWhiney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Muidasa,M, Murphy,M, Natr,L, Nakevicius,C, Neal,D, Newton,G, Olariunsaqoon,A, Pal,S, Parks,K, Naokoeleneh,O, Okwouu,G, Olariunsaqoon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Plankoch,C, Plopper,F, Polidexter,A, Popovic,D, Primus,E, Pu,L, Plazzo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R, Rellly,B, Rellly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savary,G, Scheier,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D, Speed,A, Sodergren,E, Song,X,Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczyk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstein,G, and Gibbs,R.A.

TITLE

JOURNAL

2 (bases 1 to 66877)

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (24-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 66877)

Rat Genome Sequencing Consortium

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:24942360.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GCMW

Center clone name: CH230-76N21

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 5466 bases at least Q40

Consensus quality: 58172 bases at least Q30

Consensus quality: 60570 bases at least Q20

Estimated insert size: 39548; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 38 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1213: contig of 1213 bp in length

1214 1313: gap of unknown length

1314 2354: contig of 1041 bp in length

2355 2454: gap of unknown length

2455 3485: contig of 1031 bp in length

3486 3585: gap of unknown length

3586 4610: contig of 1025 bp in length

4611 4710: gap of unknown length

4711 5813: contig of 1103 bp in length

5814 5913: gap of unknown length

5914 5977: contig of 1064 bp in length

5978 7077: gap of unknown length

7078 8546: contig of 1459 bp in length

8547 8647: gap of unknown length

8648 9573: contig of 1327 bp in length

9574 10073: gap of unknown length

10074 11267: contig of 1194 bp in length

11268 11367: gap of unknown length

11368 12462: contig of 1095 bp in length

12463 12562: gap of unknown length

12563 13862: contig of 1300 bp in length

13863 13962: gap of unknown length

13963 15078: contig of 1116 bp in length

15079 15178: gap of unknown length

15179 16297: contig of 1119 bp in length

16298 16397: gap of unknown length

16398 18180: contig of 1783 bp in length

18181 18280: gap of unknown length

18281 19887: contig of 1607 bp in length

19888 19987: gap of unknown length

19988 21410: contig of 1423 bp in length

21411 21510: gap of unknown length

21511 23019: contig of 1509 bp in length

23020 23119: gap of unknown length

23120 24271: contig of 1152 bp in length

24272 24371: gap of unknown length

24372 25380: contig of 1009 bp in length

25381 25480: gap of unknown length

25481 25847: contig of 1367 bp in length

```

source
1. 66877 /organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-76N21"
BASE COUNT 16495 a 15556 c 15614 g 15499 t 3713 others

```

Db	Accession	Version	Source	Organism	Reference Authors	Journal Title	Reference Authors
462	AC103943	190843 bp	DNA	linear	HTG 21-ANG-200		
31006	TCCGTACATGCTTTATCTCCTCGCAGAGACATTCGG	31042					
RESULT 12	AC103943/c						
LOCUS	Mus musculus clone RP23-460118		WORKING DRAFT	SEQUENCE	6 ordered		
DEFINITION	pieces						
ACCESSION	AC103943						
VERSION	AC103943.3	GI:22381785					
KEYWORDS	HTG; HTGS; PHASE2; HTGS; DRAFT; HTGS_FULLTROP.						
SOURCE	Mus musculus (house mouse)						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
AUTHORS	1 (bases 1 to 190843)						
TITLE	Biren, B., Nusbaum, C. and Lander, E.						
JOURNAL	Mus musculus, clone RP23-460118						
REFERENCE	2 (bases 1 to 190843)						
AUTHORS	Biren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barrera, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, R., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferrelle, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Hefford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Lacroque, K., Lamazeres, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhach, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Vial, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.						
TITLE	Submitted Submission						
JOURNAL	Directed Submission						
REFERENCE	Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA						
AUTHORS	3 (bases 1 to 190843)						
TITLE	Biren, B., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, R., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferrelle, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhach, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Vial, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.						
TITLE	Submitted Submission						
JOURNAL	Submitted (21-ANG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA						
COMMENT	On Aug 21, 2002 this sequence version replaced g1:19745075.						
	All repeats were identified using RepeatMasker:						
	Smith, A.F.A. & Green, P. (1996-1997)						

Db 245 CAACTGCTAACCAAGCCCTTCTCTCTGCTGAATGGGACACAGCATGCTGATGTCAC 304
LOCUS
DEFINITION Rattus norvegicus microtubule-associated proteins 1A and 1B
ACCESSION 418 GCGCATCGGAGCATCTACAGCAGAGAAAGACGAGCGCTTCTCTATATGCTCA 477
VERSION 305 ACCCATCTCTGAAGTGTACGAGAGAGAGATGAAGCGCTTCTCTATATGCTCA 364
KEYWORDS
ORGANISM Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS Brooksbank, R.A., Dixon, A.K., Lee, K. and Pinnock, R.D.
TITLE Identification and use of molecules implicated in pain
JOURNAL Patent: EP 1284297-A 72 19-FEB-2003;
WARNER-LAMBERT COMPANY (US)
FEATURES
source
1. 861
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/note="light chain 3 subunit of microtubule-associated
proteins 1A and 1B"
BASE COUNT 215 a 255 c 213 g 178 t
ORIGIN
Query Match 37.3%; Score 238.6; DB 6; Length 861;
Best Local Similarity 76.6%; Pred. No. 1.7e-30;
Matches 292; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 118 GCGGCCCCGCGGCGCGCGGATGCCCTCAAGCCGCTTCAAGCAGCGCGGAGCTT 177
Db 5 CGAGCCAGAGACCCCGCGCGCATGCGCTCCGAGAAAGCTTCAAGACGCGCGAGCTT 64
QY 178 CGCCGACCGCTGTAAAGAGTACAGAGATCCGCGACAGACCCAGCAAAATCCCGT 237
Db 65 CGAACAAAGAGTGAAGATGTCCGCTCATCCGGAGAGACAGACCCAGCAAGATCCAGT 124
QY 238 GATCATCGAGCGCTACAAAGGTGAGAAAGACAGCTGCCGCTCCGACAAAGCAAGTTT 297
Db 125 GATTATAGAGCGATACAAAGGTGAGAAAGACAGCTGCCGCTCCGACAAAGCAAGTTT 184
QY 298 GGTCCCGGACATGTCAACATGAGCGAGTGTCAAGATCATCCGGCGCGCGCTGCAGCT 357
Db 185 TGTACCTGATACCTGATATGAGCGAAGCTCATCAAGATTAATTGAGAGCGCGCTGCAGCT 244
QY 358 GAACCCACGAGCGCTTCTCTCTGCTGTAACAGACAGCATGCTGATGCTGCTAC 417
Db 245 CAATGCTAACCAAGCTTCTCTCTCTGCTGTAATGGGACAGCATGGGATGCTGCTAC 304
QY 418 GCGCATCGGAGCATCTACAGCAGAGAAAGACGAGCGCTTCTCTATATGCTCA 477
Db 305 ACCCATCTCTGAAGTGTACGAGAGAGAGATGAAGCGCTTCTCTATATGCTCA 364
QY 478 GCGCTCCAGAAACCTTCGG 498
Db 365 TGCCTCCAGAGAGAGCTTCGG 385

RESULT 15
RN005784

LOCUS
DEFINITION Rattus norvegicus microtubule-associated proteins 1A and 1B
ACCESSION 005784
VERSION 005784.1 GI:455108
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS Mann, S.S. and Hammarback, J.A.
TITLE Molecular characterization of light chain 3. A microtubule binding
JOURNAL J. Biol. Chem. 269 (15), 11492-11497 (1994)
MEDLINE 94209331
PUBMED 7908909
REFERENCE
AUTHORS Hammarback, J.A.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1994) James A. Hammarback, Neurobiology and
Anatomy, Bowman Gray School of Medicine, Medical Center Boulevard,
Winston-Salem, NC 27157-1010, USA
FEATURES
source
1. 861
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="brain"
/clone_lib="Lambda ZAP II XR (Stratagene, Inc.)"
/dev_stage="adult"
27. 455
/codon_start=1
/product="light chain 3 subunit of microtubule-associated
proteins 1A and 1B"
/protein_id="AAA20645.1"
/db_xref="GI:455109"
/translation="MPSEKFKORSFEQREVDYRLIREQPTKIPVIERKGEKOL
PVLKTKFLVDPDHVMSLIIKIRRLDNNQAFLLVNHSHSVSTPISVEYSE
RDEDFLIVMYASQETFTALAVTMSALKATATAREPCL"
845. 850
polyA_signal
polyA_site
BASE COUNT 215 a 255 c 213 g 178 t
ORIGIN
Query Match 37.3%; Score 238.6; DB 10; Length 861;
Best Local Similarity 76.6%; Pred. No. 1.7e-30;
Matches 292; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 118 GCGGCCCCGCGGCGCGCGGATGCCCTCAAGCCGCTTCAAGCAGCGCGGAGCTT 177
Db 5 CGAGCCAGAGACCCCGCGCGCATGCGCTCCGAGAAAGCTTCAAGACGCGCGAGCTT 64
QY 178 CGCCGACCGCTGTAAAGAGTACAGAGATCCGCGACAGACCCAGCAAAATCCCGT 237
Db 65 CGAACAAAGAGTGAAGATGTCCGCTCATCCGGAGAGAGACAGACCCAGCAAGATCCAGT 124
QY 238 GATCATCGAGCGCTACAAAGGTGAGAAAGACAGCTGCCGCTCCGACAAAGCAAGTTT 297
Db 125 GATTATAGAGCGATACAAAGGTGAGAAAGACAGCTGCCGCTCCGACAAAGCAAGTTT 184
QY 298 GGTCCCGGACATGTCAACATGAGCGAGTGTCAAGATCATCCGGCGCGCGCTGCAGCT 357
Db 185 TGTACCTGATACCTGATATGAGCGAAGCTCATCAAGATTAATTGAGAGCGCGCTGCAGCT 244
QY 358 GAACCCACGAGCGCTTCTCTCTGCTGTAACAGACAGCATGCTGATGCTGCTAC 417
Db 245 CAATGCTAACCAAGCTTCTCTCTCTGCTGTAATGGGACAGCATGGGATGCTGCTAC 304
QY 418 GCGCATCGGAGCATCTACAGCAGAGAAAGACGAGCGCTTCTCTATATGCTCA 477

Db 405 ACCGATCTCTGAGTGTACGAGAGCGAGAGATGAAGCGGCTTCCGTGATGTCTA 364
UY 478 CGGCTCCCGAGGAACCTTCGG 498
|||||
Db 465 TGCTCCCGAGSAGAGTTTCGG 385

Search completed: August 15, 2003, 08:56:26
Job time : 2460 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2003, 04:50:54 ; Search time 258 seconds
(without alignments)
6696.280 Million cell updates/sec

Title: US-09-904-603-2

Perfect score: 640
Sequence: 1 CTCGCCGACGCCGACGCCGCC.....GCCGCCCTACTGACAGGGCA 640

Scoring table:
IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	640	100.0	642	19	AAV43780
2	627.4	98.0	1000	24	ABQ54481
3	627	98.0	951	20	AAx20495
4	456.4	71.3	704	23	AAST3980
5	330.8	51.7	363	14	AAO61379
6	297.8	46.5	1029	24	AB199821
7	292	45.6	794	23	AAST3981
8	255.2	39.9	570	22	ABAB1070

9	255.2	39.9	570	22	AAK09363	Human brain expres
10	255.2	39.9	570	22	AAK35252	Human bone marrow
11	255.2	39.9	570	22	AA140967	Probe #9653 used 1
12	255.2	39.9	570	23	ABH4996	Human liver stifle
13	238.6	37.3	861	24	ABK61807	Kat sequence clone
14	215.4	33.7	1910	24	AB199776	Human secretory po
15	215	33.6	768	22	AA158815	Human polynucleoti
16	215	33.6	791	22	AA195937	Human neuroblastom
17	215	33.6	2250	21	AACT7815	Human cancer assoc
18	215	33.6	2309	22	AA166692	Human polynucleoti
19	215	33.6	3143	24	ABH48852	Human tumour suppr
20	214.4	33.5	2199	24	ABH52817	Genomic JNA encodi
21	212	33.1	515	22	AAH81833	Human Mpl's encodi
22	192.4	30.1	537	24	ABK63290	Kat sequence clone
23	174.2	27.2	870	21	AAH08590	Human cytoskeleton
24	163	25.5	163	22	AAH72009	Human foetal liver
25	163	25.5	163	22	AAK20434	Human brain expres
26	163	25.5	163	22	AAK46575	Human bone marrow
27	163	25.5	163	22	AA152410	Probe #21096 used
28	163	25.5	163	23	ABH46354	Human liver stifle
29	157.8	24.7	374	22	AAK53312	Human polynucleoti
30	138.6	21.7	603	22	AAK52328	Human polynucleoti
31	138.4	21.6	471	22	AAK61647	Human immature/hema
32	118.4	18.5	736	24	ABH034854	Oligonucleotide to
33	118.4	18.5	736	24	ABH034855	cDNA encoding huma
34	108.4	16.9	201	24	ABH52811	DNA encoding novel
35	106.4	16.6	990	23	AAST70567	Human secreted pro
36	99.6	15.6	535	20	AAx20427	Human breast cell
37	95.6	14.9	474	22	ABH42209	Probe #885 for gen
38	95.6	14.9	474	22	ABH52631	Human brain expres
39	95.6	14.9	474	22	ABH22419	Human bone marrow
40	95.6	14.9	474	22	AAK00892	Human liver stifle
41	95.6	14.9	474	22	AAK26347	Human polynucleoti
42	95.6	14.9	474	22	AA110980	Probe #913 for gen
43	95.6	14.9	474	22	AA132240	Probe #925 used to
44	95.6	14.9	474	22	AA100501	Probe #892 used to
45	95.6	14.9	474	23	ABH25938	Human liver single

ALIGNMENTS

RESULT 1	
AAV43780	
ID	AAV43780 standard; cDNA; 642 bp.
XX	
AC	AAV43780;
XX	
DT	20-NOV-1998 (first entry)
XX	
DE	Human microtubule associated protein hltc3 nucleotide sequence.
XX	
KW	ss: human: microtubule-associated protein; hMAP; cell proliferation;
KW	cancer.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
FT	142..507
FT	/*tag a
FT	/product= "hltc3"
PN	
XX	W09837197-A1.
PD	27-AUG-1998.
XX	
PR	23-FEB-1998; 98WD-US03744.
XX	
PR	24-FEB-1997; 97US-0805117.
XX	
PA	(INCY-) INCYTE PHARM INT.
XX	
PI	Goll SK, Hillman JL;

Sequence 1000 BP: 201 A; 337 C; 290 G; 171 T; 1 other:

Query Match 98.0%; Score 627.4; DB 24; Length 1000;

Best Local Similarity 99.7%; Pred. No. 1,5e-126; Matches 639; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 CTCCCGAGAGCCGAGCCGCTGCTCAGCGGAGAGCCCGGAGCCCTTAGCGGAGGCG 60
   |||||||
Db 6 CTCCCGAGAGCCGAGCCGCTGCTCAGCGGAGAGCCCGGAGCCCTTAGCGGAGGCG 65
QY 61 GGAG-CCCGGAGGCCCGCCAAACCGCAGACATCCCGCGCCGAGAGCCCGGCTGCG 119
   |||||||
Db 66 GGAGCCCGCGGAGGCCCGCCAAACCGCAGACATCCCGCGCCGAGAGCCCGGCTGCG 125
QY 120 CGCCCGAGCGCGCGCGCGGATGCGCTCAGACCGCGCTTTTAAGCAGTGGGCGAGCTTGG 179
   |||||||
Db 126 CGCCCGAGCGCGCGCGCGGATGCGCTCAGACCGCGCTTTTAAGCAGTGGGCGAGCTTGG 185
QY 180 CCGAGCCGCTGTAGAGAGGTACAGAGATCCCGGAGCAGCCGAGCAAAATCCGCGTGA 239
   |||||||
Db 186 CCGAGCCGCTGTAGAGAGGTACAGAGATCCCGGAGCAGCCGAGCAAAATCCGCGTGA 245
QY 240 TCATCGAGCGCTACAGAGGTGAGAGCAGCTGCCCGCTCTGAGACAGCCAGTTTGG 299
   |||||||
Db 246 TCATCGAGCGCTACAGAGGTGAGAGCAGCTGCCCGCTCTGAGACAGCCAGTTTGG 305
QY 300 TCCTCGAGCATGTACAGATGAGCGAGTTGCTCAGATCATCCGCGCGCGCTGAGCTGA 359
   |||||||
Db 306 TCCTCGAGCATGTACAGATGAGCGAGTTGCTCAGATCATCCGCGCGCGCTGAGCTGA 365
QY 360 ACCCGCAGCAGAGCCCTTCTCTGCTGTGAGACAGCAGATGATGATGCTGTCACGC 419
   |||||||
Db 366 ACCCGCAGCAGAGCCCTTCTCTGCTGTGAGACAGCAGATGATGATGCTGTCACGC 425
QY 420 CCATCGCGGAGCATTCAGACAGAGAAAGCAGAGCGCTTCTCTATATGCTACAG 479
   |||||||
Db 426 CCATCGCGGAGCATTCAGACAGAGAAAGCAGAGCGCTTCTCTATATGCTACAG 485
QY 480 CCTCCGAGGAAGCTTGGCTTCTGAGCAGCAGTAGGGGGGCTGGGCTGGAGTGGG 539
   |||||||
Db 486 CCTCCGAGGAAGCTTGGCTTCTGAGCAGCAGTAGGGGGGCTGGGCTGGAGTGGG 545
QY 540 GGGGCGCGGCTGAGGCGCTTGCACAGAGAGCTTGTGTTCTGAGCTGAGCTGCTTACCG 599
   |||||||
Db 546 GGGGCGCGGCTGAGGCGCTTGCACAGAGAGCTTGTGTTCTGAGCTGAGCTGCTTACCG 605
QY 600 TGGTGGGCTGGGCGAGCATGTGCCCGCTTACTCAGAGGCGCA 640
   |||||||
Db 606 TGGTGGGCTGGGCGAGCATGTGCCCGCTTACTCAGAGGCGCA 646

```

RESULT 3
AAAX20495
ID AAAX20495 standard; DNA: 951 BP.
AC AAAX20495;
XX
XX 04-MAY-1999 (first entry)
DE Human secreted protein gene.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW congenital disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm;
XX
XX Homo sapiens.
OS
XX
FN W9906423-A1.

```

XX 11-FEB-1999.
PD
XX 29-JUL-1998; 98W0-US15949.
PF
XX
PR 19-AUG-1997; 97US-0056740.
PR 30-JUL-1997; 97US-0054209.
PR 30-JUL-1997; 97US-0054211.
PR 30-JUL-1997; 97US-0054212.
PR 30-JUL-1997; 97US-0054213.
PR 30-JUL-1997; 97US-0054214.
PR 30-JUL-1997; 97US-0054215.
PR 30-JUL-1997; 97US-0054217.
PR 30-JUL-1997; 97US-0054218.
PR 30-JUL-1997; 97US-0054234.
PR 30-JUL-1997; 97US-0054236.
PR 18-AUG-1997; 97US-0055968.
PR 18-AUG-1997; 97US-0055969.
PR 18-AUG-1997; 97US-0055972.
PR 19-AUG-1997; 97US-0056534.
PR 19-AUG-1997; 97US-0056543.
PR 19-AUG-1997; 97US-0056554.
PR 19-AUG-1997; 97US-0056561.
PR 19-AUG-1997; 97US-0056727.
PR 19-AUG-1997; 97US-0056729.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Carter KC, Endress GA, Fan P, Feng P, Kyaw H, Lafleur LW;
PI Li Y, Moore PA, Kosen CA, Ruben SM, Shi Y, Wei Y;
PI Zeng Z;
XX WPI: 1995-153691/13.
XX P-PSDB: AAY00340.
DR
XX
XX New isolated human genes and the secreted polypeptides they encode -
PI useful for diagnosis and treatment of e.g. cancers, neurological
PI disorders, immune diseases, inflammation or blood disorders
XX
XX Disclosure: Page 249; 312pp; English.
XX
XX This sequence represents a nucleic acid molecule which encodes a
CC secreted human protein. The gene number is given in the descriptor line.
CC The gene can be used to generate fusion proteins by linking the gene
CC to a human immunoglobulin Fc portion (e.g. AAAX20403) for increasing the
CC stability of the fused protein as compared to the human protein only.
CC The invention relates to 84 novel genes and their fragments (nucleic
CC acid sequences: AAAX20412-X20499; amino acid sequences AAY00258-Y00377)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 84
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAAX20412 for described uses).
XX
XX Sequence 951 BP: 159 A; 332 C; 287 G; 170 T; 3 other:

```

Query Match 98.0%; Score 627; DB 20; Length 951;
Best Local Similarity 99.5%; Pred. No. 1,8e-126;
Matches 638; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 CTCCCGCAGATGCGAGGCGCGGTGTGAGGCGAGCGCGGAGCGCTTAGTGCAGAGCG 60
   |||||||
Db 6 CTCCCGCAGATGCGAGGCGCGGTGTGAGGCGAGCGCGGAGCGCTTAGTGCAGAGCG 65
QY 61 GGAG-CCCGGAGGCCCGCCAAACCGCAGACATCCCGCGCCGAGAGCCCGGCTGCG 119
   |||||||
Db 66 GGAGCCCGCGGAGGCCCGCCAAACCGCAGACATCCCGCGCCGAGAGCCCGGCTGCG 125
QY 120 CGCCCGAGCGCGCGCGCGGATGCGCTCAGACCGCGCTTTTAAGCAGTGGGCGAGCTTGG 179
   |||||||
Db 126 CGCCCGAGCGCGCGCGCGGATGCGCTCAGACCGCGCTTTTAAGCAGTGGGCGAGCTTGG 185

```

QY 180 CCGACGGCTGTAGAGAGTACACAGATCCGCGACGACCCAGCAAAATCCCGTGA 239
 DB 186 CCGACGGCTGTAGAGAGTACACAGATCCGCGACGACCCAGCAAAATCCCGTGA 245
 QY 240 TCATCGAGGCTTACAGAGGTGAGAGACAGTCCGCTCGGACAGACCAAGTTTGG 299
 DB 246 TCATCGAGGCTTACAGAGGTGAGAGACAGTCCGCTCGGACAGACCAAGTTTGG 305
 QY 300 TCCCGAGCATGTCAACATGAGAGAGTGGTCAATCATCCGCGCCGCTGACACTGA 359
 DB 306 TCCCGAGCATGTCAACATGAGAGAGTGGTCAATCATCCGCGCCGCTGACACTGA 365
 QY 360 ACCCCAGAGAGGCTTCTCTGCTGCTGTAACACAGACAGTGAATGTGTCCAGCG 419
 DB 366 ACCCCAGAGAGGCTTCTCTGCTGCTGTAACACAGACAGTGAATGTGTCCAGCG 425
 QY 420 CCATCGCGAGCATGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
 DB 426 CCATCGCGAGCATGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
 QY 480 CCTCCAGAGAACTTCCGCTTCTGTAGCCAGCATAGTAGGGGGCTCGCCTGGAGTGGG 539
 DB 486 CCTCCAGAGAACTTCCGCTTCTGTAGCCAGCATAGTAGGGGGCTCGCCTGGAGTGGG 545
 QY 540 GGGCCCCGGTCAAGGCGCTGCGCCAGAGAGAGTGGTCTGTAAGTGAAGTGGCTTACCG 599
 DB 546 GGGCCCCGGTCAAGGCGCTGCGCCAGAGAGTGGTCTGTAAGTGAAGTGGCTTACCG 605
 QY 600 TGGTGGGCTGGGAGGAGCATGTGCCCCCTAGTACAGAGGCA 640
 DB 606 TGGTGGGCTGGGAGGAGCATGTGCCCCCTAGTACAGAGGCA 646

RESULT 4
 AAS73980
 ID AAS73980 standard; cDNA: 704 BP.
 AC AAS73980;
 XX
 DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #9784.
 XX
 XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.

XX WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 DR P-PSDB: ABG09793.
 XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX

PS Claim 1; SEQ ID No 9784; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX

SQ Sequence 704 BP: 131 A; 263 C; 191 G; 119 T; 0 other;

Query Match 71.3%; Score 456.4; DB 23; Length 704;
 Best Local Similarity 98.8%; Pred. No. 1.3e-89;
 Matches 512; Conservative 0; Mismatches 1; Indels 5; Gaps 5;

QY 1 CTCGCCAGCCGCGAGCGCGCTGTGTCAGGCGGAGCGCGGAGCCCTTGAGCGCGAGGCC 60
 DB 15 CTCGCCAGCCGCGAGCGCGCTGTGTCAGGCGGAGCGCGGAGCCCTTGAGCGCGAGGCC 74
 QY 61 GGAG-CGCCGAGCCGCCAAGCCAGACACATCCCGCGGCCAGAGAGCCCGGCTGCG 119
 DB 75 GGAGCCCGCGGAGCCGCCAAGCCAGACACATCCCGCGGCCAGAGAGCCCGGCTGCG 134
 QY 120 GCGCCAGCGGCGCGCGCGGATGCTTCTAGAGCCGCTTTCAAGCAGCGCGGAGCTTGC 179
 DB 135 GCGCCAGCGGCGCGCGCGGATGCTTCTAGAGCCGCTTTCAAGCAGCGCGGAGCTTGC 194
 QY 180 CCGA-CCGCTGTAGAGAGTACAGAGATCCG-CGACGACGACCCAGCAAAATCCCGGT 237
 DB 195 CCGACCCGCTGTAGAGAGTACAGAGATCCGCGACGACGACCCAGCAAAATCCCGGT 254
 QY 238 -GATCATGAGCGCTTCAAGAGGTGAGAGACAGTCCCGCTGTGACAGAGCAAGTTT 296
 DB 255 GATCATGAGCGCTTCAAGAGGTGAGAGACAGTCCCGCTGTGACAGAGCAAGTTT 314
 QY 297 TGGTCCCGGAGCATGTCAACATGAGAGAGTGGTCAATATCTCCGGCGCGCTGACAGC 356
 DB 315 TGGTCCCGGAGCATGTCAACATGAGAGAGTGGTCAATATCTCCGGCGCGCTGACAGC 374
 QY 357 TGAACCCGAGCGAGGCTTCTTCTGCTGCTGTAACACAGCAGCATGTGTGAGTGTG-TCC 415
 DB 375 TGAACCCGAGCGAGGCTTCTTCTGCTGCTGTAACACAGCAGCATGTGTGAGTGTG-TCC 434
 QY 416 AGGCCATCGCGGAGATCTAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 475
 DB 435 AGGCCATCGCGGAGATCTAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
 QY 476 TAGGCTCCAGAGAACTTGGGCTTGGAGCGAGCAG 513
 DB 495 TAGGCTCCAGAGAACTTGGGCTTGGAGCGAGCAG 532

RESULT 5
 AAO61379
 ID AAO61379 standard; DNA: 363 BP.
 XX
 AC AAO61379;
 XX

DT 25-MAR-2003 (updated)
 DT 16-MAR-1994 (first entry)
 XX

DE Human brain Expressed Sequence Tag EST01371.
 XX
 XX Gene transcription product: genetic markers; tagging: in vivo;
 KW transcription; mapping; locations; chromosomes; chromosomal; ss.
 XX
 XX Homo sapiens.
 XX
 XX W09316178-A2.
 XX
 XX 19-AUG-1993.
 XX
 XX 12-FEB-1993: 93WO-0501294.
 XX
 XX 12-FEB-1992: 92US-0837195.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX
 XX Adams MD, Moreno RF, Ventler CJ;
 XX
 XX WPI: 1993-272882/34.
 XX
 XX Enriched oligonucleotides and corresp. sequences - used as
 PT markers for human genes transcribed in-vivo, facilitate tagging
 PT of most human genes
 XX
 XX Example 4: Page 482: 500pp: English.
 XX
 XX The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC for human genes transcribed in vivo. They can be used to facilitate
 CC tagging of most human genes, for mapping locations of expressed genes
 CC on chromosomes, for individual or forensic identification, for mapping
 CC locations of disease-associated genes, for identification of tissue
 CC type, and for prepn. of antisense sequences, probes and constructs.
 CC EST01371 has a "excellent" coding probability as evaluated using the
 CC coding-region prediction program CRM. See also AA059041-061440.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC
 XX Sequence 363 BP: 84 A; 108 C; 99 G; 66 T; 6 other:
 SO
 Query Match 51.7%; Score 330.8; DB 14; Length 363;
 Best Local Similarity 97.3%; Pred. No. 1.8e-62;
 Matches 354: Conservative 0; Mismatches 8; Indels 2; Gaps 2;
 QY 153 GGCCTTTACAGCAGCGCGGAGCTTCGCCGACCGCTGTGAAGAGGTACACGACATCCCG 212
 DB 1 GGCCTTTACAGCAGCGCGGAGCTTCGCCGACCGCTGTGAAGAGGTACACGACATCCCG 59
 QY 213 ACCAGACCCCGACAAATCCGGTGCATGACGCGGTACAAAGGTGAGAAGCAGCTGC 272
 DB 60 ACCAGACCCCGACAAATCCCGGTCATGACGCGGTACAAAGGTGAGAAGCAGCTGC 119
 QY 273 CCGTCTGGACAGACACAGTCTTGTGTCGCCGACCATGTCAACATGAGCGAGTTGCTCA 332
 DB 120 CCGTCTGGACAGACACAGTCTTGTGTCGCCGACCATGTCAACATGAGCGAGTTGCTCA 179
 QY 333 AGATCATCGGGCGCGCTGAGTGAACCCACGAGCGCTTCTTCTGCTGGTGAAC 392
 DB 180 AGATCATCGGGCGCGCTGAGTGAACCCACGAGCGCTTCTTCTGCTGGTGAAC 239
 QY 393 AGCAGCAGTGTGTGTCACAGCCCATGCGGACATCTACAGAGAGAGAAAGAG 452
 DB 240 AGCAGCAGTGTGTGTCACAGCCCATGCGGACATCTACAGAGAGAGAAAGAG 299
 QY 453 AGAGAGCTTCTCTATATGTGTACAGCGCTCCAGAAACCTTCGGC-TTCTGAGCCAGC 511
 DB 300 AGGAGCGCTTCTCTATATGTGTACAGCGCTCCAGAAACCTTCGGCTTCTGAGNCAGC 359
 QY 512 AGTA 515
 DB 360 AGTA 363

RESULT 6
 ABL99821
 ID ABL99821 standard; cDNA: 1029 BP.
 XX
 XX ABL99821:
 AC
 XX
 XX 03-OCT-2002 (first entry)
 DT
 XX
 XX Human secretory polynucleotide (sptm) 76.
 DE
 XX
 XX Human; ss: gene; secretory protein; secretory polynucleotides; SPIM;
 KW SPTM-related disease; somatic gene therapy; demiline gene therapy;
 KW severe combined immunodeficiency; intracellular parasite protection;
 KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
 KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
 KW motor neuron disorder; demyelinating disease; multiple sclerosis;
 KW meningitis; abscess; prion diseases; cerebral palsy;
 KW neuroskeletal disorder; peripheral nervous system disorder;
 KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
 KW mental disorder; Loureite's syndrome.
 XX
 XX Homo sapiens.
 OS
 XX
 XX W0200220756-A2.
 XX
 XX 14-MAR-2002.
 PD
 XX
 XX 30-AUG-2001: 2001WO-US27297.
 PF
 XX
 XX 05-SEP-2000: 2000US-229747P.
 PR
 XX 05-SEP-2000: 2000US-229748P.
 PR
 XX 05-SEP-2000: 2000US-229749P.
 PR
 XX 05-SEP-2000: 2000US-229750P.
 PR
 XX 05-SEP-2000: 2000US-229751P.
 PR
 XX 05-SEP-2000: 2000US-230016P.
 PR
 XX 05-SEP-2000: 2000US-230583P.
 PR
 XX 05-SEP-2000: 2000US-230514P.
 PR
 XX 05-SEP-2000: 2000US-230515P.
 PR
 XX 05-SEP-2000: 2000US-230516P.
 PR
 XX 05-SEP-2000: 2000US-230517P.
 PR
 XX 05-SEP-2000: 2000US-230518P.
 PR
 XX 05-SEP-2000: 2000US-230519P.
 PR
 XX 05-SEP-2000: 2000US-230555P.
 PR
 XX 05-SEP-2000: 2000US-230596P.
 PR
 XX 05-SEP-2000: 2000US-230597P.
 PR
 XX 05-SEP-2000: 2000US-230599P.
 PR
 XX 05-SEP-2000: 2000US-230610P.
 PR
 XX 05-SEP-2000: 2000US-230644P.
 PR
 XX 05-SEP-2000: 2000US-230865P.
 PR
 XX 05-SEP-2000: 2000US-230988P.
 PR
 XX 05-SEP-2000: 2000US-230989P.
 PR
 XX 05-SEP-2000: 2000US-230990P.
 PR
 XX 05-SEP-2000: 2000US-230996P.
 PR
 XX 05-SEP-2000: 2000US-230897P.
 PR
 XX 05-SEP-2000: 2000US-230951P.
 PR
 XX 05-SEP-2000: 2000US-231163P.
 PR
 XX 05-SEP-2000: 2000US-231832P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX
 XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
 PI Jones AL, Yu JY, Wright RJ, Gietzen U, Liu TF, Yap PE, Dahl CK;
 PI Momiyama MC, Bradley DL, Kohatgi SD, Harris B, Roseberry AM;
 PI Gerstin EH, Petalita CH, David MB, Panzer SR, Flores V, Datto A;
 PI Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR;
 XX
 XX WPI: 2002-315658/35.
 DR
 XX P-PSDB: ABB97824.
 PT
 XX Polynucleotide sequences encoding human secretory proteins useful for
 PT gene therapy of e.g. genetic deficiency disorders, cancers, and
 PT diseases caused by intracellular parasites -

XX
PS
xy

Claim 1; Page 296; 585pp; English

CC The invention comprises the amino acid and coding sequences of human
CC secretory (SPM), proteins. The SPM DNA and amino acid sequences are
CC useful for treating a disease or condition associated with the expression
CC of functional SPM. The SPM DNA sequences are useful for somatic or
CC germline gene therapy to correct a genetic deficiency (e.g. severe
CC combined immunodeficiency). The SPM DNA sequences are also useful in
CC providing protection against intracellular parasites (e.g. fungal
CC parasites and protozoan parasites). The SPM DNA and protein sequences
CC are also useful for diagnosing cell proliferative disorders, cancer,
CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
CC sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
CC neuroskeletal disorders, peripheral nervous system disorders,
CC dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
CC disorders (e.g. Tourette's syndrome). cDNA sequences ABL97946 - ABL99929
CC represent human secretory polynucleotides of the invention.

Sequence 1029 BP; 172 A; 318 C; 354 G; 183 T; 2 other;

Query Match	46.58;	Score 297.8;	DB 24;	Length 1029;
Best Local Similarity	00.39;	Score 10.5;	DB 1;	Length 1029;

Matches	299;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

QY	340	CGGGGCGCCCTGACAGCTGAACCCAGCCAGGCGCTTCTCTCGTGGGAAACGACACAG	399
Db	554	CAGGGCGCCCTGACAGCTGAACCCAGCCAGGCGCTTCTCTCGTGGGAAACGACACAG	613
QY	400	CATGGAGATGTGTCCAGCCCATATCGCGACATCTACGACGAGAGAAGAAGACGAGACGG	459
Db	614	CATGGAGATGTGTCCAGCCCATATCGCGACATCTACGACGAGAGAAGAAGACGAGACGG	673
QY	460	CTTCCCTATATAGTGTCTACGCCCTCCACAGAAACCTCGCGCTTCTGAGACGACAGTACGGG	519
Db	674	CTTCCCTATATAGTGTCTACGCCCTCCACAGAAACCTCGCGCTTCTGAGACGACAGTACGGG	733
QY	520	GGCTCGGCGCTGGGAATCGGGGGGGCCCGCGTAAAGGCCCTGCCACAGAGACTTCTGGTTCTT	579
Db	734	GGCTCGGCGCTGGGAATCGGGGGGGCCCGCGTAAAGGCCCTGCCACAGAGACTTCTGGTTCTT	793
QY	580	GAACCTAGATGCTCTTACCGTGTGGGTGGGCGAGCATGTGCCCCCTAGTACAGAGGGC	639
Db	794	GAACCTAGATGCTCTTACCGTGTGGGTGGGCGAGCATGTGCCCCCTAGTACAGAGGGC	853
QY	640	A 640	
Db	854	A 854	

RESULT 7
AAS73981
ID AAS73981 standard; cDNA; 794 BP.

DT 13-FEB-2002 (first entry)
 YY

DNA encoding novel human diagnostic protein #9785.

human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167
YY

PA (HYSE-) HYSEQ INC.
VY

PI Drmanac RT, Liu C, Tang YT;
v v

DR WPI; 2001-639362/73.
D-P-DEDB; APC00704

XX
XX
New 301 and 302.

PT new isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 1; SEQ ID No 9785; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. A6564197-A6594564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/publ/published_pat_sequences.

50 Sequence 794 BP; 151 A; 253 C; 231 G; 159 T; 0 other,

Query Match	Score	DB	Length
45.68	292	23	794

Matches	401;	Conservative	0;	Mismatches	20;	Indels	17;	Gaps	8.
---------	------	--------------	----	------------	-----	--------	-----	------	----

207 TCCGCGACCAGCACCCAGCAAAATCCCGGTGATCATCGAGCGCTACAAGGTTGAGAAGC 266

Db	ITCCGGGACGACGACCCACGACAAATCCCGGTGATCATGAGGGCTACAAAGGTTGAGAAC	60
OY	267 ACCTGCCCCCTCTGTGACAAGACCAAGTTTGTGTCCGGACCATGTCAACTATGAGCGAGT	327
Db	61 AACTCCCGCTCTGTGACAAGACCAAGTTTGTGTCCGGACCATGTCAACTATGAGCGAGT	120
OY	327 TGTGTAAGATCATCTCCGGGCGCGCTCGACAGCTGAACCCGACGAGCCCTTATCCGCTGG	387
Db	121 TGTGTAAGATCATCTCCGGGCGCGCTCGACAGCTGAACCCGACGAGCCCTTATCCGCTGG	187
OY	387 TGAACCCAGCACGATGTTGAGTGTGTCCAGGCCCATCGCGACATCTACGACGACGAGA	447
Db	181 TGAACCCAGCACGATGTTGAGTGTGTCCAGGCCCATCGCGACATCTACGACGACGAGA	241
OY	447 AAGAC-GAGGACGGCTTCT-CATATGCTTACGGCTCC--AGGAACCTTGGGCTT	501
Db	241 AAGAGGGAGGACGGCTTCTCTATATATGTCTACGGCTCCAGGGAACCTTGGGCTT	301
OY	502 CTGACCC-AGCAAGTAGGGGGGCTCGGCT-GGAGATGGGGGGGCCCGGTACGGCTT-G	558
Db	301 CTGACCCAGGAGTAGGGGGGCTCGGCTTGGGGAGATGGGGGGGCCCGGTACGGCTTGG	360
OY	559 CCCAGAGAGCTTCTGGTCTCTGAA-----CTGACCTGGCTCTACCGTGGTGGGCTGG	610
Db	361 CCCAGAGAGCTCTGGTCTCTGAA-----CTGACCTGGCTCTTAAACCGGTGGTGGGCTGG	420
OY	611 GCA-GGCATGTCGCCCCC 627	

Db 431 CTCTCTATATGCTCTACGGCTCCAGAAACCTTCGGCTTCTGAGCCAGACTAGGG 490
 QY 520 GGCTCGGCTGGAGTGGGGGGCCCGGTCAGGCGCTTCCAGAGACTTCTGTTCT 579
 Db 491 GGCTCGGCTGGAGTGGGGGGCCCGGTCAGGCGCTTCCAGAGACTTCTGTTCT 550
 QY 580 GAAGTGAAGTGGCTCTACGG 599
 Db 551 GAAGTGAAGTGGCTCTACGG 570

RESULT 12

ABK34996
 ID ABS34996 standard; DNA: 570 BP.
 XX
 AC ABS34996;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver single exon probe, SEQ ID No 9986.
 XX
 KW Human: single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200157273-A2.
 XX
 PD 05-AUG-2001.
 XX
 PE 30-JAN-2001: 2001WO-0500664.
 XX
 PR 04-FEB-2000: 2000US-0180312.
 PR 26-MAY-2000: 2000US-0207456.
 PR 30-JUN-2000: 2000US-0608408.
 PR 03-AUG-2000: 2000US-0632366.
 PR 21-SEP-2000: 2000US-0234687.
 PR 27-SEP-2000: 2000US-0236359.
 PR 04-OCT-2000: 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver.
 XX
 PS Claim 1; SEQ ID No 9986; 658bp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SEN) (1) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (1) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABS25011-ABS51005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pub/published/pct_sequences.
 XX
 SO Sequence 570 BP: 97 A; 158 C; 213 G; 102 T; 0 other;

Query Match 39.94; Score 255.2; DB 23; Length 570;
 Best Local Similarity 98.88; Pred. No. 4.3e-46;
 Matches 257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 340 CCGGCGCGGCTGGAGTGGAGCCGACGACGACGCTTCTCTCTGTTGAGATCAACAAG 399
 Db 311 CAGGCGCGGCTGGAGTGGAGCCGACGACGACGCTTCTCTCTGTTGAGATCAACAAG 370
 QY 400 CATGGTGAAGTGTTCACAGCCCAAGCGGATATCTAGAGAGAGAGAAAGAGAGAGAG 459
 Db 371 CATGGTGAAGTGTTCACAGCCCAAGCGGATATCTAGAGAGAGAGAAAGAGAGAGAG 430
 QY 460 CTTCCTCTATATGCTCTACGGCTCCAGAAACCTTCGGCTTCTGAGCCAGACTAGGG 519
 Db 431 CTCTCTATATGCTCTACGGCTCCAGAAACCTTCGGCTTCTGAGCCAGACTAGGG 490
 QY 520 GGCTCGGCTGGAGTGGGGGGCCCGGTCAGGCGCTTCCAGAGACTTCTGTTCT 579
 Db 491 GGCTCGGCTGGAGTGGGGGGCCCGGTCAGGCGCTTCCAGAGACTTCTGTTCT 550
 QY 580 GAAGTGAAGTGGCTCTACGG 599
 Db 551 GAAGTGAAGTGGCTCTACGG 570

RESULT 13

ABK63807
 ID ABK63807 standard; cDNA: 861 BP.
 XX
 AC ABK63807;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Rat sequence differentially expressed in response to a hepatotoxin #1714.
 XX
 KW Rat; ss; hepatotoxicity; expressed sequence tag; EST; drug screening;
 KW differential expression; centrilobular necrosis; steatosis.
 XX
 OS Rattus norvegicus.
 XX
 PN W0200210453-A2.
 XX
 PD 07-FEB-2002.
 XX
 PE 30-JUL-2001: 2001WO-0523872.
 XX
 PR 31-JUL-2000: 2000US-222040P.
 PR 02-NOV-2000: 2000US-244880P.
 PR 11-MAY-2001: 2001US-290029P.
 PR 15-MAY-2001: 2001US-290645P.
 PR 22-MAY-2001: 2001US-292336P.
 PR 06-JUN-2001: 2001US-295798P.
 PR 13-JUN-2001: 2001US-297457P.
 PR 19-JUN-2001: 2001US-298884P.
 PR 09-JUL-2001: 2001US-303459P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter MW, Johnson KR, Castle AL, Flashoff MR;
 XX
 DR WPI: 2002-241625/29.
 XX
 PT Predicting toxic effects of compounds or the progression of these toxic
 PT effects by determining the changes in gene expression in tissues of
 PT cells exposed to the toxin and comparing these to gene expression in
 PT unexposed tissues or cells.
 XX
 PS Claim 1; Seq ID No 1714; 239bp; English.
 XX
 CC The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods or predicting at least one toxic
 CC effect of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2003, 08:06:39 : Search time 74 Seconds
(without alignments)
3817.368 Million cell updates/sec

Title: US-09-904-603-2

Perfect score: 640

Sequence: 1 CTCGCCGAGCGCGAGCGCC.....GCCGCCCTAGTCAGAGGSCA 640

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:
1: /cqn2-6/ptodata/1/lna/5A_COMB.seq:*
2: /cqn2-6/ptodata/1/lna/5B_COMB.seq:*
3: /cqn2-6/ptodata/1/lna/6A_COMB.seq:*
4: /cqn2-6/ptodata/1/lna/6B_COMB.seq:*
5: /cqn2-6/ptodata/1/lna/PCTUS_COMB.seq:*
6: /cqn2-6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	640	100.0	640	2	US-08-805-117-2
2	640	100.0	640	3	US-09-199-838-2
3	215	33.6	768	4	US-09-620-312D-707
4	214.4	33.5	2199	4	US-09-708-725A-3
5	108.4	16.9	201	4	US-09-708-725A-1
6	53.6	8.4	939	4	US-09-252-991A-4937
7	53.6	8.4	1317	4	US-09-252-991A-4950
8	53.6	8.4	1932	4	US-09-252-991A-4950
9	49.8	7.8	1764	4	US-09-252-991A-1709
10	49	7.7	2038	4	US-09-620-312D-597
11	47.2	7.4	519	4	US-09-702-705-78
12	47.2	7.4	519	4	US-09-736-457-78
13	47	7.3	1950	3	US-08-911-851-14
14	47	7.3	1950	3	US-09-479-409-14
15	47	7.3	1950	4	US-09-479-453-14
16	47	7.3	17612	3	US-08-911-853-29
17	47	7.3	17612	3	US-09-479-409-29
18	47	7.3	17612	4	US-09-479-453-29
19	45.6	7.1	477	2	US-08-665-647-12
20	45.6	7.1	990	4	US-09-252-991A-11517
21	45.6	7.1	2166	4	US-09-252-991A-11722
22	45.4	7.1	906	4	US-09-252-991A-1857
23	44.8	7.0	1575	4	US-09-252-991A-11212
24	44.8	7.0	1620	4	US-09-252-991A-11146
25	44.8	7.0	2472	4	US-09-252-991A-11043
26	44.8	7.0	45272	1	US-08-614-770A-1
27	44.2	6.9	1995	1	US-08-425-069-3

28	44.2	6.9	1995	2	US-08-117-844B-3	Sequence 3, Appl1
29	43.4	6.8	3318	4	US-09-593-589-3	Sequence 3, Appl1
30	42.8	6.7	1659	4	US-09-084-351-3	Sequence 3, Appl1
31	42.8	6.7	1659	3	US-09-084-352-3	Sequence 3, Appl1
32	42.8	6.7	3946	3	US-09-084-351-1	Sequence 1, Appl1
33	42.8	6.7	3946	3	US-09-084-352-1	Sequence 1, Appl1
34	42.2	6.6	2433	4	US-09-620-312D-830	Sequence 830, App
35	42.2	6.6	2598	4	US-09-417-197-110	Sequence 110, App
36	42.2	6.6	2615	4	US-09-417-197-108	Sequence 108, App
37	41.8	6.5	777	4	US-09-252-991A-6204	Sequence 6204, App
38	41.8	6.5	1836	4	US-09-252-991A-6042	Sequence 6042, App
39	41.8	6.5	1911	4	US-09-252-991A-6258	Sequence 6258, App
40	41.8	6.5	35081	2	US-08-752-760A-1	Sequence 1, Appl1
41	41.6	6.5	942	4	US-09-252-991A-5840	Sequence 5840, App
42	41.6	6.5	1503	4	US-09-464-702A-33	Sequence 33, Appl
43	41.6	6.5	1503	4	US-09-469-135-33	Sequence 33, Appl
44	41.6	6.5	2007	4	US-09-252-991A-5770	Sequence 5770, App
45	41.6	6.5	3627	4	US-09-792-594-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-805-117-2
Sequence 2, Application US/08805117
Patent No. 5955312
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,117
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET INFORMATION: IF-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 640 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYRNTO3
CLONE: 1441378
US-08-805-117-2
Query Match 100.0% Score 640; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 1,96-134;
Matches 640; Conservative 0; Mismatches 0; Gaps 0;

QY 1 CTCGCCAGCCGCGAGCCGCGTCTCAGCGCGAGCCCGGAGCCCTTGAGCCGAGCGGC 60
1 CTCGCCAGCCGCGAGCCGCGTCTCAGCGCGAGCCCGGAGCCCTTGAGCCGAGCGGC 60
QY 61 GGAGCCCCGGAGCCCCCAACCGCGAGACATCCCCGGCGCCAGAGCCCCGGCTGGCG 120
61 GGAGCCCCGGAGCCCCCAACCGCGAGACATCCCCGGCGCCAGAGCCCCGGCTGGCG 120
QY 121 GCCAGCGCGGCGCGCGATGCTCCTCAGACCGCGCTTTCAAGACGCGGAGGCTTGGC 180
121 GCCAGCGCGGCGCGCGATGCTCCTCAGACCGCGCTTTCAAGACGCGGAGGCTTGGC 180
QY 181 CGACCGCTGTAAAGAGGTACAGACATCCGCGACAGACCCAGCAAAATCCCGGTAT 240
181 CGACCGCTGTAAAGAGGTACAGACATCCGCGACAGACCCAGCAAAATCCCGGTAT 240
QY 241 CATGAGGGCTACAGGGGTGAGAACAGCTGCCCTCTCGCAAAAGACCAAGTTTGGT 300
241 CATGAGGGCTACAGGGGTGAGAACAGCTGCCCTCTCGCAAAAGACCAAGTTTGGT 300
QY 301 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
301 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
QY 361 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
361 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
QY 421 CATGCGGACATCTCAGAGGAGAGAAAGAGAGAGCGCTTCTATATAGTGTACGC 480
421 CATGCGGACATCTCAGAGGAGAGAAAGAGAGAGCGCTTCTATATAGTGTACGC 480
QY 481 CTCGAGGAAACCTTCGCTCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
481 CTCGAGGAAACCTTCGCTCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GCGCCCGGTACAGGGGTGAGAACAGCTGCGCTCTGGAAGTGTGCTTACCGT 600
541 GCGCCCGGTACAGGGGTGAGAACAGCTGCGCTCTGGAAGTGTGCTTACCGT 600
QY 601 GGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGCA 640
601 GGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGCA 640

RESULT 2
US-09-199-838-2
; Sequence 2, Application US/09199838
; Patent No. 6280733
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL MICROTUBULE-ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jincyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/199,838
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,117

QY 1 CTCGCCAGCCGCGAGCCGCGTCTCAGCGCGAGCCCGGAGCCCTTGAGCGCGAGCGGC 60
1 CTCGCCAGCCGCGAGCCGCGTCTCAGCGCGAGCCCGGAGCCCTTGAGCGCGAGCGGC 60
QY 61 GGAGCCCCGGAGCCCCCAACCGCGAGACATCCCCGGCGCCAGAGCCCCGGCTGGCG 120
61 GGAGCCCCGGAGCCCCCAACCGCGAGACATCCCCGGCGCCAGAGCCCCGGCTGGCG 120
QY 121 GCCAGCGCGGCGCGCGATGCTCCTCAGACCGCGCTTTCAAGACGCGGAGGCTTGGC 180
121 GCCAGCGCGGCGCGCGATGCTCCTCAGACCGCGCTTTCAAGACGCGGAGGCTTGGC 180
QY 181 CGACCGCTGTAAAGAGGTACAGACATCCGCGACAGACCCAGCAAAATCCCGGTAT 240
181 CGACCGCTGTAAAGAGGTACAGACATCCGCGACAGACCCAGCAAAATCCCGGTAT 240
QY 241 CATGAGGGCTACAGGGGTGAGAACAGCTGCCCTCTCGCAAAAGACCAAGTTTGGT 300
241 CATGAGGGCTACAGGGGTGAGAACAGCTGCCCTCTCGCAAAAGACCAAGTTTGGT 300
QY 301 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
301 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
QY 361 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
361 CCCGACCATGTCAACATGAGCGAGTGTGTCAGATCCGCGCGCGCTCAGCTGAA 360
QY 421 CATGCGGACATCTCAGAGGAGAGAAAGAGAGAGCGCTTCTATATAGTGTACGC 480
421 CATGCGGACATCTCAGAGGAGAGAAAGAGAGAGCGCTTCTATATAGTGTACGC 480
QY 481 CTCGAGGAAACCTTCGCTCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
481 CTCGAGGAAACCTTCGCTCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GCGCCCGGTACAGGGGTGAGAACAGCTGCGCTCTGGAAGTGTGCTTACCGT 600
541 GCGCCCGGTACAGGGGTGAGAACAGCTGCGCTCTGGAAGTGTGCTTACCGT 600
QY 601 GGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGCA 640
601 GGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGCA 640

Query Match 100.0%; Score 640; DB 3; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.9e-134; Mismatches 0; Gaps 0;
Matches 640; Conservative 0; Indels 0;

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 640 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYRNOT03
CLONE: 141378
US-09-199-838-2

RESULT 3
US-09-620-312D-707
; Sequence 707, Application US/09620312D
; Patent No. 6569662


```

: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Ren, Feiyan
: APPLICANT: Chen, Rui-hong
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wehrman, Tom
: APPLICANT: Xue, Aidong J.
: APPLICANT: Yang, Yonghong
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Zhou, Ping
: APPLICANT: Ma, Yundong
: APPLICANT: Wang, Dunrui
: APPLICANT: Wang, Zhimei
: APPLICANT: John Tillinghast
: APPLICANT: Drmanac, Radoje T.
: TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
: TITLE OF INVENTION: Polypeptides
: FILE REFERENCE: 784CIP2B
: CURRENT APPLICATION NUMBER: US/09/620,312D
: CURRENT FILING DATE: 2000-07-19
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: 09/488,725
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 1105
: SOFTWARE: Pf_genes Version 1.0
: SEQ ID NO 707
: LENGTH: 768
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (109)..(486)
: US-09-620-312D-707

Query Match      33.6%: Score 215; DB 4; Length 768;
Best Local Similarity 74.1%: Pred. No. 1.5e-39;
Matches 272; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 132 CCGCGGGATGCCCTCAGACGCGGCCCTTCAAGCAGCGCGAGCTTGGCTGACCGCTGTA 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 101 CCGCCACCATGCCCTCGAGAGACACTTCAAGCAGCGCGGACCTTGCACAAAGAGTAG 160
QY 192 AGGAGTACACAGATCCGGGACGACGACCCGACCAAAATCCGGGTATCATCGAGCGCT 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 161 AAGATGTCCGACTTATTCGAGAGCGACATCCACCAAAATCCCGGTATATAGACGAT 220
QY 252 ACAAGGCTGAGAACGACCTGCCCTGCTGGACAGACCAAGTTTGTGCGCGGACGATG 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 221 ACAAGGCTGAGAACGACCTTCTTCTTGATTAACAAGTTCTTGTACTTACCATG 280
QY 312 TCAACATGAGCGAGTGTGTCAGATCATCCGCGCGCTGCGCATGAAACCCACGCGAG 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 281 TCAACATGAGTGAAGCTCATAGATTAATTAAGAAAGCGGTTACAGCTCATGTATCAG 340
QY 372 CCTTCTCTCTGCTGTGTAACGACGACGATGTGTGTCCAGCGCCATCGCGGACA 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 341 CCTTCTCTCTGCTGTGTAACGACGACGATGTGTGTCCAGCGCCATTCAGAGG 400
QY 432 TCTACGAGCTGAGAGAGAGAGAGAGGAGGCTTCTCTATGCTATGAGCGTCCGAGAAA 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 401 TGTATGAGAGTGAAGAAAGATGAGATGATGATTCGTGATCATGTCTATGCTCCAGGAGA 460
QY 492 CCTTGGG 498
    |||||
DB 461 CGTTCGG 467
```

RESULT 4
US-09-708-725A-3

```

: Sequence 3, Application US/09708725A
: Patent No. 6489456
: GENERAL INFORMATION:
: APPLICANT: LADUNGA et al.
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: C1000842
: CURRENT APPLICATION NUMBER: US/09/708,725A
: CURRENT FILING DATE: 2000-11-09
: PRIOR APPLICATION NUMBER: 60/243,428
: PRIOR FILING DATE: 2000-10-27
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 2199
: TYPE: DNA
: ORGANISM: Human
: US-09-708-725A-3

Query Match      33.5%: Score 214.4; DB 4; Length 2199;
Best Local Similarity 68.5%: Pred. No. 2.5e-39;
Matches 296; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 67 CCGGAGTCCGCCAAGTCGATACACATCCCGCGCGCCGAGAGCCCGGCTGACCGCCAG 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 750 CCGGAGTACGGCGGCGACGCTGAGAGCGCGCGCGGACCGCTCGGCTGTGCTGCGCGCG 809
QY 127 CCGGAGTCCGCCAAGTCGATACACATCCCGCGCGCCGAGAGCTTTCGCGGACCG 186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 810 AGATCCGTCACACGATAGTGGTGGAGAAATCTTCAGTACGCGGCGCAATCTGAAATAAAG 869
QY 187 CTGTAAAGGAGGTACAGAGATCCGCGGACGACATCCGCAAAATCCGCTGATGATGCA 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 870 AGTAAAGATGTCCGACTTATTCGAGAGCGACATCCCAAAATCCCGCTGATATAGTA 929
QY 247 GCGCTACAGGCTGAGAGAGTACGCTGCGCTGCGTACAGACCAAGTTTGTGCTCGCGA 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 930 ACGATACAGGCTGAGAGAGTACGCTGCGCTGCGTATGATTAACCAAGTTCTGTGATCGTGA 989
QY 307 CCGATGCAACATGAGCGAGTTGTCAAGATATCCGCGCGCGCTGCAACATGAAATCGTAA 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 990 CCGATGCAACATGAGTGTGATGATCAAGATTAATTAAGAAAGCGGCTTACAGCTCAATGCTAA 1049
QY 367 GCGCGCTCTCTCTGCTGTGATGACGACGACGATGATGATGATGATGATGATGATGAT 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1050 TCAGGCGCTCTCTCTGCTGTGATGACGACGACGATGATGATGATGATGATGATGATGAT 1109
QY 427 GCGATCTACGAGAGAGAAACGAGAGAGGCTTCTCTATATGATGATGATGATGATGATGAT 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1110 AAGAGTGTATGAGAGTGAAGAAAGATGAGATGATGATGATGATGATGATGATGATGATGAT 1159
QY 487 GGAACCTTGGG 498
    ||||| |||||
DB 1170 GGAAGAGCTTGGG 1181
```

RESULT 5
US-09-708-725A-1
: Sequence 1, Application US/09708725A
: Patent No. 6489456
: GENERAL INFORMATION:
: APPLICANT: LADUNGA et al.
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: C1000842
: CURRENT APPLICATION NUMBER: US/09/708,725A
: CURRENT FILING DATE: 2000-11-09
: PRIOR APPLICATION NUMBER: 60/243,428
: PRIOR FILING DATE: 2000-10-27
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1
LENGTH: 201
TYPE: DNA
ORGANISM: Human
US-09-708-725A-1

Query Match 16.9%; Score 108.4; DB 4; Length 201;
Best Local Similarity 74.7%; Pred. No. 7.6e-16;
Matches 136; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 317 ATGACGAGTGTGTAAGATCATCCGGCGCCCTGCAGCTGAACCCACGAGGCTTC 376
DB 1 ATGAGTGAAGCTCATCAATATTAAGAGCGCTTACAGCTCAAGCTAATGAGGCTTC 60
QY 377 TTCTGCTGTGAACACACAGCATGTGTGTCACGCGCCATCGGAGCATCTAC 436
DB 61 TTCTGCTGTGTGAACAGACACAGCATGTGTGTCACGCGCTTCACACCAATCTCAAGGTGTAT 120
QY 437 GAGCAGGAAAGACGAGAGCGGCTTCTCTATATATGTTACGCTCCAGGAAACCTTC 496
DB 121 GAGAGTGAAGAAAGATGATGATGATCTCTGATACATGCTGTGCTCCAGAGACGTTTC 180
QY 497 GG 498
DB 181 GG 182

RESULT 6

US-09-252-991A-4937
Sequence 4937, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 4937
LENGTH: 939
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4937

Query Match 8.4%; Score 53.6; DB 4; Length 939;
Best Local Similarity 45.3%; Pred. No. 0.0017;
Matches 194; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY 57 GCGGAGAGCCCGGAGCCCGCAAAACCGAGACATCCCGCGCCCGCCAGAGCCCGGCT 116
DB 12 GCGCCAGAGTGGCATCCCGTACTGAGCTTCCGANTCCCGGCGCCGCGACAGCGCG 71
QY 117 GCGGCGCCAGCGGCGCCCGCGATGCCCTCAGACCGCGCTTTCAAGCAGCGCGAGCT 176
DB 72 AGGGATCCGCCAGAGGCGATGGCGCTTTCGACGACCTCAACACACCGCGGATCCGCA 131
QY 177 TCGCGGACCGCTGTAAAGAGGTACAGCAGATCCCGGACGACACCCCGAGCAAAATCCCG 236
DB 132 TCGGCTTCGCGCCCGCGAGCTCTTATAGGTGAGCGACGACACTGGAGAGATCTCG 191
QY 237 TGATCATGAGCGCTTACAGGAGTGAAGAGCAGCTGCCGCTCTGAGACAAAGCTTTT 296
DB 192 TGTCTACCGAGAACTGAGCGCCAGCATCCAGATGACGCTCCAGACCGCTTCGAGG 251
QY 297 TGTCTCCGAGCATGTCAACATGAGCGAGTGTGTCAGATCATCGGCGCGGCTCGAGC 356
DB 252 TGGAGCAGGCGCAGGAGCGCAAGCGCGCGCTGCGCTGCGCTGCGCTCGGCGC 311

QY 357 TGAACCCAGCAGCGCTTCTTCTGCTGTGTAACACGACGACATGGTGAAGTGTCCA 416
DB 312 TGTCTGCGCCAGCGCTTCCAGGCGGTGTCACATGACCCAGTGGACGACGATCGGCA 371
QY 417 GCGCCATGCGGAGCATCTACGAGCAGGAGAAAGACGAGCGGCTTCTCTATATGCTCT 476
DB 372 TGTCTGTGAAACCAACAGCTGTGATTCACCTGCCGGAATTCACCTCAAGCTTGCGCA 431

QY 477 ACCGCTCC 484
DB 432 GCGGCTTC 439

RESULT 7

US-09-252-991A-4950
Sequence 4950, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 4950
LENGTH: 1317
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4950

Query Match 8.4%; Score 53.6; DB 4; Length 1317;
Best Local Similarity 45.3%; Pred. No. 0.0018;
Matches 194; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY 57 GCGGAGAGCCCGGAGCCCGCAAAACCGAGACATCCCGCGCCCGCCAGAGCCCGGCT 116
DB 428 GCGCCAGAGTGGCATCCCGTACTGAGCTTCCGANTCCCGGCGCCGCGACAGCGCG 487
QY 117 GCGGCGCCAGCGGCGCCCGCGATGCCCTCAGACCGCGCTTTCAAGCAGCGCGAGCT 176
DB 488 AGGGATCCGCCAGAGGCGATGGCGCTTTCGACGACCTCAACACACCGCGGATCCGCA 547
QY 177 TCGCGGACCGCTGTAAAGAGGTACAGCAGATCCCGGACGACACCCCGAGCAAAATCCCG 236
DB 548 TCGGCTTCGCGCCCGCGAGCTCTTATAGGTGAGCGAGACACTGGAGAGATCTCG 607
QY 237 TGATCATGAGCGCTTACAGGAGTGAAGAGCAGCTGCCGCTCTGAGACAAAGCTTTT 296
DB 608 TGTCTACCGAAGAACTGAGCGCCAGCATCCAGATGACGCTCCAGAGACCGCTTCGAG 667
QY 297 TGTCTCCGAGCATGTCAACATGAGCGAGTGTGTCAGATCATCGGCGCGGCTTCGAGC 356
DB 668 TGGAGCAGGCGCAGGAGCGCAAGCGCGCGCTTGGCGGCTGTGACCGCTTCGCGC 727
QY 357 TGAACCCAGCAGCGCTTCTTCTGCTGTGTAACACGACGACAGCATGTTGATGTCCA 416
DB 728 TGTCTGCGCCCGCGAGCTTTCAGGCGGTGTCACATGACCCAGGTGAGAGAGCACTTGGCA 787
QY 417 CGCCATGCGGAGCATGTACGAGCAGGAGAAAGACGAGCGGCTTCTCTATATGCTCT 476
DB 788 TGTCTGTGAAACCAACAGCTGTGATTCACCTGCCGGAATTCACCTCAAGCTTGCGCA 847
QY 477 ACCGCTCC 484
DB 848 GCGGCTTC 855

RESULT 8

US-09-252-991A-4998/C
: Sequence 4998, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 4998
: LENGTH: 1932
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4998

Query Match 8.4%; Score 53.6; DB 4; Length 1932;
Best Local Similarity 45.3%; Pred. No. 0.0019;
Matches 194; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY 57 GCGCGAGACCCCGAGCCCAAAACCGGAGACATCCCGCGCGCCGAGAGCCCGCCCT 116
DB 1025 GCGCCCAAGTGGCGATCCCGGTACTGCGACTTCGATCCCGCGCGCCGAGAGCCCG 966
QY 117 GCGCGCCGAGCCGCGCGCCGCGCGATGCCCTGAGACCGCGCCCTTCAAGCAGCGCGCGAGCT 176
DB 965 AGCGGATCCGCGAGCGCGCATGGCGCTGTGAGCAGACTCAAGCAGCAGCCCGCGCATGCCA 906
QY 177 TCGCGGACCGCTGTAGAGAGGTACAGCAGATCCCGGACAGCAGCCCGCAAAATCCCGG 236
DB 905 TCGCGCTCGCCCGCCCGCCCTCTTATACGATGAGGAGGAGCAGCAAGCTGGAGCAGATCCCTGG 846
QY 217 TGATATGAGCGCGCTACAAAGGAGTGAAGAGCGTCCCTCGGAGCAAGACCAAGTTT 296
DB 845 TGCACAGCGAGGAGTCCAGCCCGCAGCATCCAGATGCGAGTCCAGCAGACCCGCTTCGAGG 786
QY 297 TGTCCCGGAGCATGTCAACATGAACGAGTTGCTCAAGATCATCCGCGCGCCCTCGCAGC 356
DB 785 TGGAGCAAGCGCCAGCGAGCGCAACGCGCGCCGCTGCGCCCTGCAACCGCTCGCGCC 726
QY 357 TGAACCGCAGCGAGCGCTTCTTCTGCTGCTGTAACACAGCAGATGCTGAGTGTGTCCA 416
DB 725 TGTGCGCGCCAGCGCTTCCAGCGCGGTGCACATGACCGCAGTGGACGACGACGACCTGGCGGA 666
QY 417 CGCCCATCGCGAGCATCTACGAGCAGAGAGAAAGAGAGAGCGGCTTCTATATATGCTCT 476
DB 665 TGTGTGTGAAGCAACGAGCTGCGGTGATGATCAGTCCGCGGAATCCAACTCAAGCTGGCCA 606
QY 477 AGCGCTTC 484
DB 605 GCGGCTTC 598

RESULT 9
US-09-252-991A-1709
: Sequence 1709, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27

: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 1709
: LENGTH: 1764
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1709

Query Match 7.8%; Score 49.8; DB 4; Length 1764;
Best Local Similarity 44.3%; Pred. No. 0.013;
Matches 245; Conservative 0; Mismatches 307; Indels 1; Gaps 1;

QY 17 GCGCGTCTCAGCGCGAGAGGCGGAGCGCTTGAAGCGCGAGCGCGCGAGCGCGCGAGCGCG 76
DB 681 GCGCGAGCGAGTGTCTGATGAGCGAGCTGCTGCGCGGCTGTGCGCGCGCGCGCGCGCGCG 740
QY 77 CAACCGCAGACACATCCCTCGCGCGCGAGAGCGCGCGCTGTGCGCGCGCGCGCGCGCGCG 145
DB 741 TCGACCGCGAGCG 809
QY 136 CGCGATGCGCGTGCAGCGCGCGCTTGAAGCAGCGCGCGAGCTTGCAGCGCGCGCGCGCGCG 195
DB 801 CGCGATGCGCGAGCGCGCGCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860
QY 196 GGTACAGCAGATCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 255
DB 861 CGAGCGGTACCGCGCTGAGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 920
QY 256 GAGTGAAGAGTGTGCG 415
DB 921 GCG 980
QY 316 CAAGCGCGAGTGTGCG 475
DB 981 GAGCGGTATGCTGCG 1040
QY 376 CTGCTGCTGCGCGAGCG 435
DB 1041 GATCG 1106
QY 436 CGAGCAGCAGAAAGCG 495
DB 1101 GCTCGATGTGAGCG 1160
QY 496 CGCGCTTGTGAGCG 555
DB 1161 CGACCTGACTCG 1220
QY 556 CTGCGCAGAGAGC 568
DB 1221 CTGCGCGCGCAGC 1233

RESULT 10
US-09-620-312D-597
: Sequence 597, Application US/09620312D
: Patent No. 6569662
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Ren, Feiyao
: APPLICANT: Chen, Rui-hong
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wehrman, Tom
: APPLICANT: Xue, Aidong J.
: APPLICANT: Yang, Yonghong
: APPLICANT: Wang, Jian-kui
: APPLICANT: Zhou, Ping
: APPLICANT: Ma, Yundong
: APPLICANT: Wang, Duanfui
: APPLICANT: Wang, Zhiwei
: APPLICANT: John Tiliudast

DB 826 CGCCTACGCGACGAGCCCAAGACCTTGCCCGAGACCTTGCGGACATCTCCGAGGG 885
QY 242 ATCGAGCGCTTACAGAGGTGAGAGCAGTCCCGTCTGAGACAGCAAGTTTGTGTC 301
DB 886 TTGAAGACCGCGAGGTGTGCGGACGCGGACAGAGCCGAGAAACATCTCGATC 945
QY 302 CCGGACATGTAAATGAGGAGTGTGCAAGATCATCGGCGCCCTGACGTGANC 361
DB 946 CCGGCGGAGAGGCTCAATCCCTGCTGCTGCGGATCCGACACCCCTGCGGACC 1005
QY 362 CCCAGCAGGCGCTTCTGCTGCTGAGAACAGCAGCATGTGTGATGTGTCACGCGCC 421
DB 1006 CTCGAGGAAATGTGTCGCAACCTCGACGTGCGCGCGCCGACAGTGTGTCGAGGCGGC 1065
QY 422 ATCGGCGACATCTACAGAGAGAGAAAGAGAGAGGCTTCTCTATATGCTTACGCC 481
DB 1066 ATCGTGAATCTCCGGGACATCAAGCAGCGCCCTCGGCGGTGACATGGGCGGTGATGCC 1125
QY 482 TCCCAGGAAACCTTGCGCTTCTGAG 506
DB 1126 CGCGGCGGACCGCGGCTCGCG 1150

RESULT 15
US-09-479-453-14
Sequence 14, Application US/0949453
Patent No. 6313283
GENERAL INFORMATION:
APPLICANT: Geriltse, Gijdsbert
APPLICANT: Quax, Wilhelmus J.
TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
TITLE OF INVENTION: EXPRESSION LEVELS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Genencor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/479,453
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,853
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Debra J
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC361-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7620
TELEFAX: 650-845-6504
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-479-453-14

Query Match 7.3%, Score 47; DB 4; Length 1950;

Best Local Similarity 45.6%, Pred. No. 0.057;

Matches 203; Conservative 0; Mismatches 240; Indels 2; Gaps 1;

QY 64 GCCCGGAGCCCGCAACCGAGACACATCCCGCGCCGAGAGCCCGGCTGGCG--G 121
DB 706 GCCGACGCGCGCACCAACCGCGCTGATCTCTCGCGCGCGCGCGCGCGCGCGCTG 765

QY 122 CCCAGCGGCGCCCGCGATGCGCTTACAGCCGCGCTTTCAGACAGCGGAGCTTGGCC 181
DB 766 GCCAAGCTGCGCGCGCTCGGTGACATCCCGACACCCGCTTGGCCAAATCCGGGGTAATT 825
QY 182 GACCGCTGTAAGGAGGTACAGAGATCCGCGACAGCAGCCGAGCAAAATCCCGGTGATC 241
DB 826 CGCCTACGCGACAGCGAGCGCCAAAGAGCCTGGCCGAGACCTTGGGACATCTCCGAGGG 885
QY 242 ATCGAGCGCTTACAGAGGTGAGAGAGAGTGGCCGCTCTGAGACAAAGAGTTTGTGTC 301
DB 886 TTGAAGACCGCGAGAGGTGTGCGAGAGCGCCAGCAGCAGCAACCCGAGAACTCTGATC 945
QY 302 CCGGACATGTCAACATGAGAGAGTGTGCAAGATCATCCGCGCGCCCTGACGTGANC 361
DB 946 CCGGCGGAGAGAGCTCAATCCCTGTGCTGCTGCGGATCCGAGACACGCTGGCGACC 1005
QY 362 CCCAGCAGGCGCTTCTGCTGCTGAGAACAGCAGCATGTGTGTCACAGGCC 421
DB 1006 CTCGAGGAAATGTGTCGCAACCTCGACGTGCGCGCGCCGACAGTGTGTCGAGGCGGC 1065
QY 422 ATCGGCGACATCTACAGAGAGAGAAAGAGAGAGGCTTCTCTATATGCTTACGCC 481
DB 1066 ATCGTGAATCTCCGGGACATCAAGCAGCGCCCTCGGCGGTGACATGGGCGGTGATGCC 1125
QY 482 TCCCAGGAAACCTTGCGCTTCTGAG 506
DB 1126 CGCGGCGGACCGCGGCTCGCG 1150

Search completed: August 15, 2003, 09:32:03
Job time : 76 secs


```
OY 299 GTCCGACCATGTCACATGAGCGAGTGTGTCAGATATCCGGCGCCCTGCAGCTG 358
    |||||||
DB 333 GTCCGACCATGTCACATGAGCGAGTGTGTCAGATATCCGGCGCCCTGCAGCTG 392
OY 359 AACCCAGCAGCGCTTCTCTGCTGTGTAACACGACGACATGATGTGTGTCACG 418
    |||||||
DB 393 AACCCAGCAGCGCTTCTCTGCTGTGTAACACGACGACATGATGTGTGTCACG 452
OY 419 CCATGCGGACATATGACAGAGAGAAAGACGAGCGGCTTCTCTATATGTGTAC 478
    |||||||
DB 453 CCATGCGGACATATGACAGAGAGAAAGACGAGCGGCTTCTCTATATGTGTAC 512
OY 479 GCTTCCAGGAAACCTTCTGCTTCTGAGCCAGACAGTAGGGGGCTCGGCTGAGATCG 538
    |||||||
DB 513 GCTTCCAGGAAACCTTCTGCTTCTGAGCCAGACAGTAGGGGGCTCGGCTGAGATCG 572
OY 539 GGGGCGCCGCTGAGGCGCTTCCGCGACAGAGCTTCTGCTTCTGATGAGCTGCTAC 598
    |||||||
DB 573 GGGGCGCCGCTGAGGCGCTTCCGCGACAGAGCTTCTGCTTCTGATGAGCTGCTAC 632
OY 599 GTGTGGGCTGGGCGAGCATGTGCCCCCTAGTACAGAGGCA 640
    |||||||
DB 633 GTGTGGGCTGGGCGAGCATGTGCCCCCTAGTACAGAGGCA 674
```

RESULT 2

US-10-837-707
Sequence 707, Application US/10105837

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
APPLICANT: Dimanac, Radote T.
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2BDIVA
CURRENT APPLICATION NUMBER: US/10/105,837
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO 707
LENGTH: 768
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (109)..(486)
US-10-837-707

Query Match 33.6%, Score 215, DB 6, Length 768;
Best Local Similarity 74.1%, Pred. No. 8e-44;
Matches 272: Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
OY 132 CCGCGGCGATGCTTCAAGACCGGCGGAGCTTCCGCGACCGCTGTA 191
    |||||||
DB 101 CCGCGGCGATGCTTCAAGACCGGCGGAGCTTCCGCGACCGCTGTA 160
OY 192 AGAGGTACAGCAATCCGCGACGACGCCAGCAAAATCCGGTATCATCGAGCGT 251
    |||||||
DB 161 AAGATGTCGACTTATTCGAGAGAGCATCAACCAAAATCCGGTATTAAGAGCAT 220
OY 252 ACAAGGTGAGAGCAGCTGCGCTCCGAGAGAGCAAGCAAAATTTTGGTCCGGAGCAT 311
    |||||||
DB 221 ACAAGGTGAGAGCAGCTGCTTCTGATTAAGCAAAATTTCTTACTGACCATG 280
```

```
OY 312 TCAACATGAGCGAGTGTGTCAGATCATCCGGCGCCCTGCAGCTGAGACCCAGCAGG 371
    |||||||
DB 281 TCAACATGAGCGAGTGTGTCAGATCATCCGGCGCCCTGCAGCTGAGATGCTAATG 340
OY 372 CCTTCTCTGCTGTGTAACACGACGACATGATGTGTGTCACGCGCCATCGCGACA 431
    |||||||
DB 341 CCTTCTCTGCTGTGTAACACGACGACATGATGTGTGTCACGCGCCATCGTAGAG 400
OY 432 TCTACGACGAGAGAAAGACGAGAGCGGCTTCTCTATATGTGTACGCTCCAGAAA 491
    |||||||
DB 401 TGTATGAGAGTCAAGAAATGATGATGATTCCTGTATATGCTCTCCAGAGAGA 460
OY 492 CCTTCGG 498
    |||||||
DB 461 GCTTCGG 467
```

RESULT 3

US-10-286-897-1019
Sequence 1019, Application US/10286897

GENERAL INFORMATION:

APPLICANT: Hyseq Inc
TITLE OF INVENTION: Novel Nucleic Acid and polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/10/286,897
CURRENT FILING DATE: 2002-11-01
PRIOR APPLICATION NUMBER: US/09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US/09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US/09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US/09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US/09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US/09/662,191
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: US/09/693,036
PRIOR FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US/09/727,344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7143
SOFTWARE: PL_FL_genes_b Versions 1.0
SEQ ID NO 1019
LENGTH: 768
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (109)..(486)
US-10-286-897-1019

Query Match 33.6%, Score 215, DB 6, Length 768;
Best Local Similarity 74.1%, Pred. No. 8e-44;
Matches 272: Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
OY 132 CCGCGGCGATGCTTCAAGACCGGCGGAGCTTCCGCGACCGCTGTA 191
    |||||||
DB 101 CCGCGGCGATGCTTCAAGACCGGCGGAGCTTCCGCGACCGCTGTA 160
OY 192 AGAGGTACAGCAATCCGCGACGACGCCAGCAAAATCCGGTATCATCGAGCGT 251
    |||||||
DB 161 AAGATGTCGACTTATTCGAGAGAGCATCAACCAAAATCCGGTATTAAGAGCAT 220
OY 252 ACAAGGTGAGAGCAGCTGCGCTCCGAGAGAGCAAGCAAAATTTTGGTCCGGAGCAT 311
    |||||||
DB 221 ACAAGGTGAGAGCAGCTGCTTCTGATTAAGCAAAATTTCTTACTGACCATG 280
OY 312 TCAACATGAGCGAGTGTGTCAGATCATCCGGCGCCCTGCAGCTGAGACCCAGCAGG 371
    |||||||
DB 281 TCAACATGAGCGAGTGTGTCAGATCATCCGGCGCCCTGCAGCTGAGATGCTAATG 340
```


OY	372	CCCTTTCTCTGTTGGTGAACCGACACGCTATGGTGTGTCCACGCCCATTCGGGACA	431
Db	341	CCCTTTCTCTGTTGGTGAACCGACACGCTATGGTGTGTCCACGCCCATTCGAGAG	400
OY	432	TCATACGACAGAGAAACGACGAGCGCTTCTCTATATGTGTACGGCTCCACAGAAA	491
Db	401	TGATAGACAGTGCAGAAAGTGAAGATGATTCCTGTACATGGTCTATGGCTCCACAGAGA	460
OY	492	CCTTGCG	498
Db	461	CGTTGCG	467

```

RESULT 4
US-10-258-898A-1019
: Sequence 1019, Application US/10258898A
: GENERAL INFORMATION:
: APPLICANT: Hyseq Inc
: TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
: FILE REFERENCE: 784FLPCT
: CURRENT APPLICATION NUMBER: US/10/258,898A
: CURRENT FILING DATE: 2002-10-29
: PRIOR APPLICATION NUMBER: US/09/488,725
: PRIOR FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: US09/552,317
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: US09/598,042
: PRIOR FILING DATE: 2000-06-20
: PRIOR APPLICATION NUMBER: US09/620,312
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: US09/653,450
: PRIOR FILING DATE: 2000-08-31
: PRIOR APPLICATION NUMBER: US09/662,191
: PRIOR FILING DATE: 2000-09-14
: PRIOR APPLICATION NUMBER: US09/693,036
: PRIOR FILING DATE: 2000-10-19
: PRIOR APPLICATION NUMBER: US09/727,344
: PRIOR FILING DATE: 2000-11-29
: NUMBER OF SEQ ID NOS: 7143
: SOFTWARE: pc_Fl_genes_b Versions 1.0
: SEQ ID NO 1019
: LENGTH: 768
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURES:
: NAME/KEY: CDS
: LOCATION: (109)..(486)
: US-10-258-898A-1019

```

Query Match	33.6%	Score 215	DB 6	Length 766
Best Local Similarity	74.1%	Pred. No. 86-44		
Matches	272	Conservative 0	Mismatches 95	Indels 0
		Gaps 0		
QY	132	CCGCGCGATCCCTCGAGACCGGCGCTTTCAAGCAGCGGCGAGTTCGCGACCGCTGTA	191	
Db	101	CCCGCACCATCCCGTCGGAGAAACCTTCAAGCACGCGCGCACTCTGCACCAAGAGTAG	160	
QY	192	AGGAGGTACACAGATCCGCGACCGACGCCAGCAAAATCCGGGTGATCATGAGCGGT	251	
Db	161	AAGATGTCCGACTATTTCGACAGCAGCATCCCAACCAAAATCCGGGTGATTAATAGAACAT	220	
QY	252	ACAAGGGTGACAAACAGCTGCGCCGCTCTCGACACAGCAAGTTTGGTCCCGACCATG	311	
Db	221	ACAAGGGTGACAAACAGCTTCTCTGTCGGATAAACCAAGTTCCTTGTACTCGACATG	280	
QY	312	TCAACATGAGCAGATTGGTTCAAAGATCATCCGGCGCGCTCGACGTGAACCCACGACAG	371	
Db	281	TCACATGAGTGAGGCTCATCAAGATTAATTAGAAAGCGCGCTACAGCTCAATGCTAATCAG	340	
QY	372	CGTCTCTCTCTGCTGTAACACGACAGACAGATGATGATGTGTCCAGGCCCATCGCGACA	431	
Db	341	CGTCTCTCTCTGCTGTAACACGACAGACAGATGATGATGATGTGTCCACACCATCAATCTCAGAG	400	

OY 432 CTGATGACAGTAAAGAACAAGACGGCTTCCTCTAATGTCAAGCGCTCCGAGAAA 491
| | | | | | | | | | | | | | | | | | | | |
Db 401 TGTATGAGAGTGAAGAACATCAGATGCATTCTGTACTGGGTATGCGCTGCCAGAGA 460

OY 492 CCTTCGG 498
| | | | |
Db 461 CGTTCCG 467

```

RESULT 5
US-10-286-897-4591/c
; Sequence 4591, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPc1
; CURRENT APPLICATION NUMBER: US/10/286,897
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4591
; LENGTH: 2309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-286-897-4591

```

Query Match	33.6%	Score 215	DB 6	Length 2309
Best Local Similarity	74.1%	Pred. No. 1,1e-43		
Matches 272	Conservative 0	Mismatches 95	Indels 0	Gaps 0
QY	132	CCCGCGGCGATGCCCTCAGACCGGCGT111CAACGACGCGCGGAGCTTGGCGGCGT3151A	191	
Db	2225	CCGCGACCGATGCGCGTGGSAAGAGTCTTCAGAGCACGCGCGCAGCTTCGATCAACAGATAG	2166	
QY	192	AGGAGGTACAGCAGATCCGGGACGACGCCGAGTAAATCCCGGTGATCATCGAGCGT1	251	
Db	2165	AAGATGTGCGACTATTGAGAGGACGACGTCCAACCCAAATCCCGGTGATTAATAGAACAT	2104	
QY	252	ACAGAGGTGAGACAGCTGCCGCGCTGGGCAACGACGAGTTTGGTTCCTGATGCTAG1	311	
Db	2105	ACAAGGGTGAACAGCAGCTTCTGTTCGTGGATAAACAAAGTTCTTGTACTGACGTA1G	2044	
QY	312	TCAACATGAGCGAGTTGGTCAAGATCATCGGCGCGCGCTGCAGCTGAACGCCCAATGTA3	371	
Db	2045	TCAACATGAGTGAAGCTCTATCAAGATAATATGAAGGCGCTTACAGCTCAATGTATATCA3	1986	
QY	372	CGTCTTCTGCTGCTGGTGAACGACGACGATGAGTGTGTCCAGCGCGCATATCGCGATGA	431	
Db	1985	CGTCTTCTGCTGTGGTGAACGGAACACGATGGTCAAGGCTTCCACACGCAATATCTAGAG3	1922	
QY	432	TCTAGAGCAGAGGAGAAAGATGAGGACCGGCTTCTCTATATGTGTCTACGCTGCGTAA	491	
Db	1925	TGTATGAGAGTGAAGAAAGATGAAGATGAGATTCGTGTACATGTGTATGCTCTGCAAGGA	1866	
QY	492	CGTTGCG498		
Db	1865	CGTTGCG1859		

```
RESULT 6
US-10-258-898A-4591/c
; Sequence 4591, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pl_FL_genes_b Versions 1.0
; SEQ ID NO 4591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-898A-4591

Query Match
Best Local Similarity 33.6%; Score 215; DB 6; Length 2309;
Matches 272; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 132 CCGCGCGAGCCCTTCAGACGCGGCTTTCAGACGCGGAGCTTCGCCGACGCGCTTA 191
DB 2225 CCGCGACCATGCGCTCGGAGAGAGACCTTCAGACGCGCGGACCTTCGACAAAGAGTAG 2166
QY 192 AGGAGCTACAGCATCCGCGACGACACCCAGCAAAATCCGGTGATCATGAGCGCT 251
DB 2165 AAGAGTCCGACTTATTCGAGAGCGACATCCAAACCAATCCGGTGATTAAGAACAT 2106
QY 252 ACAAGGTGAGAGAGCGTGCCTGCGGACCAAGCTTTGGCCCGGACCGAG 311
DB 2105 ACAAGGTGAGAGAGCGCTTCCTGCTGCTGATTAACAAGTTCTTGTACTGACCATG 2046
QY 312 TCAACATGAGGAGTGTGTCAGATCATCCGCGCGCGCTGAGCTGAAACCCAGCGAG 371
DB 2045 TCAACATGAGGAGTGTGTCAGATTAATTAAGAGCGCTTACAGCTCATGTATTCAGG 1986
QY 372 CCTTCTCTGCTGCTGTAACGACGACAGCATGCTGATGCTGCAACGCCCATCGCGACA 431
DB 1985 CCTTCTCTGCTGCTGTAACGACGACAGCATGCTGATGCTGCAACCAATCTCAGAGG 1926
QY 432 TCTAGACAGGAGAAAGACGAGCGGCTTCCTATATGCTTACGCCCTCCAGGAAA 491
DB 1925 TGTATGAGAGTGAGAAAGATGAGATTCCTCTATACATGCTATCCCTCCAGGAGA 1866
QY 492 CCTTCGG 498
DB 1865 CGTTCGG 1859
```

```
RESULT 7
US-10-293-244-2841
; Sequence 2841, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
```

```
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 2841
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-293-244-2841

Query Match
Best Local Similarity 24.7%; Score 157.8; DB 6; Length 374;
Matches 210; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 206 ATCCGCGACCGACCCCAAGCAAAATCCCGGTGATCATGAGCGCTACAGGCTGACAG 265
DB 6 ATCCGCGCAAAAGTTCCCAACAAATCCCGGTGAGTAGGAGCGCTACCCAGGAGACG 65
QY 266 CAGCTGCCCCCTTCGACAGACCAAGTTTGGTCCCGGACCAATGTCATCATGAGGAG 325
DB 66 TTCTGCCCCCGCTGACAAACCAAGTTCTGCTGCTGCGGACGAGCTACCATGACCCAG 125
QY 326 TTGTCAGATCATTCGCGCGCGCTGACGTGAACCCACGACGCGCTTCTTCTGCTG 385
DB 126 TTCTCAGCATCATTCGCGCGCGCATGCTGCTGAGAGCGAGCAAGCTTTTACTTCTG 185
QY 386 GTGAACCAAGCATGCTGAGTGTGTCACAGCCCATCGCGGACATCTAGACAGAGAG 445
DB 186 GTGAACCAAGAGCGCTGCTGAGCATGAGCGCAACCATGAGAGATCTACAGAGACTAC 245
QY 446 AAAGCAGAGAGCGCTTCTCTATATGCTTACGCTCCAGGAAACCTTCGCGCTTC 502
DB 246 AAGATGAGGATGCTTCTGCTGATGACCTACGCTCCAGGAGACTTTGGCTGC 302
```

```
RESULT 8
US-09-654-936A-107
; Sequence 107, Application US/09654936A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
```

```
FILE REFERENCE: 787C1P2C
CURRENT APPLICATION NUMBER: US/09/654,936A
CURRENT FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 164
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO: 107
LENGTH: 603
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (120)..(527)
US-09-654-936A:107
```

```
Query Match      21.7%: Score 138.6; DB 5; Length 603;
Best Local Similarity 66.7%: Pred. No. 5e-25;
Matches 198: Conservative 0; Mismatches 99; Indels 0; Gaps 0;
```

```
QY 140 ATGCCCTAGACGGCGCTTTCAAGCAGCGCGGCGAGCTTCGCCGACCGCTGTAGAGAGTA 199
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 138 ATCCCAAGCGTCAAGACCCCTTCAAGCAGAGGAAAGCTTTGCAATCAGACAAAGAGAGTT 197
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 200 CAGCAGATCCGCGACACAGACCCCGCAAAATCCCGGTATCATCGACCGCTACAAAGGT 259
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 198 GCTGGAATCCGGCGAAAGTTCGCCCAACAAATCCCGGTGTAGTGAAGGCTACACCCAGG 257
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 240 GAGAACGACGCTCCCGCTCTTGACAAAGCAAGTTTGTGTCGCCGAGCATGTCAACATG 319
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 258 GAGACGTTCTCTGCCCGCTTGACAAAGCAAGTTCTGTCGCCGAGAGCTGACCAATG 317
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 320 AGCGAGTTGCTCAATATATCGCGCGCGCCCTGCACTGTAACCCAGCGAGGCTTTCTTC 379
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 318 ACCCGAGTTCTCTGAGATATCGCGAGCGCGCATGCTCTGAGAGCCAGCGAAGCTTTTAC 377
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 380 CTGCTGTGTAACGACAGCAGCATGTGTGTCCAGCGCGCATGCTGAGCTAG 436
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 378 TTGCTGTGTAACACAGAGGCTGTGTCAGCATGAGCCGACAGCATGTGACATCTAC 434
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

RESULT 9
US-10-293-244-873

Sequence 873, Application US/10293244
GENERAL INFORMATION:

APPLICANT: Hysq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom

```
SEQ ID NO 873
LENGTH: 603
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (120)..(527)
US-10-293-244-873
```

```
Query Match      21.7%: Score 148.6; DB 6; Length 603;
Best Local Similarity 66.7%: Pred. No. 5e-25;
Matches 198: Conservative 0; Mismatches 99; Indels 0; Gaps 0;
```

```
QY 140 ATGCCCTAGACGGCGCTTTCAAGCAGCGCGGCGAGCTTCGCCGACCGCTGTAGAGAGTA 199
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 138 ATCCCAAGCGTCAAGACCCCTTCAAGCAGAGGAAAGCTTTGCAATCAGACAAAGAGAGTT 197
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 200 CAGCAGATCCGCGACACAGACCCCGCAAAATCCCGGTATCATCGACCGCTACAAAGGT 259
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 198 GCTGGAATCCGGCGAAAGTTCGCCCAACAAATCCCGGTGTAGTGAAGGCTACACCCAGG 257
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 260 GAGAACGACGCTCCCGCTCTTGACAAAGCAAGTTTGTGTCGCCGAGCATGTCAACATG 319
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 258 GAGACGTTCTCTGCCCGCTTGACAAAGCAAGTTCTGTCGCCGAGAGCTGACCAATG 317
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 320 AGCGAGTTGCTCAATATATCGCGCGCGCTGCACTGTAACCCAGCGAGGCTTTCTTC 379
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 318 ACCCGAGTTCTCTGAGATATCGCGAGCGCGCATGCTCTGAGAGCCAGCGAAGCTTTTAC 377
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 380 CTGCTGTGTAACGACAGCAGCATGTGTGTCCAGCGCGCATGCTGAGCTAG 436
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 378 TTGCTGTGTAACACAGAGGCTGTGTCAGCATGAGCCGACAGCATGTGACATCTAC 434
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

RESULT 10
US-09-820-649-94

Sequence 94, Application US/09820649
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: 83 Human Secreted Proteins
FILE REFERENCE: P2012P1
CURRENT APPLICATION NUMBER: US/09/820,649
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/572,557
PRIOR FILING DATE: 1999-01-26
PRIOR APPLICATION NUMBER: 09/594,949
PRIOR FILING DATE: 1998-07-29
PRIOR APPLICATION NUMBER: 60/054,212
PRIOR FILING DATE: 1997-07-30
PRIOR APPLICATION NUMBER: 60/054,209
PRIOR FILING DATE: 1997-07-30
PRIOR APPLICATION NUMBER: 60/054,234
PRIOR FILING DATE: 1997-07-30
PRIOR APPLICATION NUMBER: 60/054,218
PRIOR FILING DATE: 1997-07-30
PRIOR APPLICATION NUMBER: 60/054,214
PRIOR FILING DATE: 1997-07-30
PRIOR APPLICATION NUMBER: 60/054,236
PRIOR FILING DATE: 1997-07-30
PRIOR APPLICATION NUMBER: 60/054,215
PRIOR FILING DATE: 1997-07-30
PRIOR APPLICATION NUMBER: 60/054,211
PRIOR FILING DATE: 1997-07-30
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 353
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 94
LENGTH: 535
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (529)

OTHER INFORMATION: n equals a,t,g, or c
US-09-820-649-94

Query Match 15.6%; Score 99.6; DB 5; Length 535;
Best Local Similarity 97.1%; Pred. No. 2e-15;
Matches 99; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 539 GGGGCCCCGGTACGAGCCCTGCCAGAGACCTTCTGTTCTGTAAGTACCTGCTTACC 598
DB 60 SGRGCCCGGCTGAGCCCTGCCAGAGACCTGTTCTGTTCTGTAAGTACCTGCTTACC 119
QY 599 GTGGTGGGCTGGGAGGAGTGGCCCCCTTACTAGAGAGGCA 640
DB 120 GTGGTGGGCTGGGAGGAGTGGCCCCCTTACTAGAGAGGCA 161

RESULT 11
US-10-626-717-7205
Sequence 7205, Application US/10626717

GENERAL INFORMATION:
APPLICANT: Andersen, Scott E.
APPLICANT: Byrum, Joseph R.
APPLICANT: De La Pena, Robert C.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
TITLE OF INVENTION: PLANTS
FILE REFERENCE: 38-21(15878)D
CURRENT APPLICATION NUMBER: US/10/626,717
CURRENT FILING DATE: 2003-07-25
PRIOR APPLICATION NUMBER: US 10/304,123
PRIOR FILING DATE: 2002-11-26
PRIOR APPLICATION NUMBER: US 09/594,596
PRIOR FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 10952
SEQ ID NO 7205
LENGTH: 448
TYPE: DNA
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: uc-tsrow189015e11b1
US-10-626-717-7205

Query Match 11.0%; Score 70.4; DB 6; Length 448;
Best Local Similarity 54.8%; Pred. No. 2.9e-08;
Matches 184; Conservative 0; Mismatches 146; Indels 6; Gaps 2;

QY 136 CTTTCAACGAGCGGCGAGCTTCCGCGACCGCTGTAAAGAGGTACAGAGATCCCGACC 215
DB 119 CGTTCAAGCTGGAGCAGCCGCTCGAGAGAGGAGCGAGGCTTAACGCAATAGAGAGA 178
QY 216 AGCAGCCGAGCAATCCCGGTGATCATCGAGCGCTACAGAGGTGAGAGCAGCTGCCG 275
DB 179 AGTACTGTGACAGAAATTCCTGTGATCGTTGAG--AAAGCTGGGAGAGGTGATATCTCTG 235
QY 276 TCTTGAGACAGCAAGTTTGTGTCGCCGACCATGTCAACATGAGCGAGTGGTCAAGA 335
DB 236 ACATTGACAAGAAAGATACCTGCTCCCGCCGACCTTACGTTGACAGTTCGTATG 295
QY 336 TCATCCGCGCGCTCGAGCTGAAGCCCGAGCGAGCTTCTTCTGCTGTGTAACCAAC 395
DB 296 TTGTCGGAGAGCGGATCAAGCTCAAGTGTGAGAAAGGAGCATCTTCACTTTGTGA--AGA 352
QY 396 ACAGCATGTAGTGTGTCACGCCATCCGCGAGCATCTAGACAGAGAGAGAGAGAG 455
DB 333 ACACCTCTTCAACGAGCTGCGCTGATGTCTGCTTACGAGAGAAAGAGAGAGAG 412
QY 456 ACAGCTCTTATATGTTAGCTACGCTCCAGAGAA 491
DB 413 ACAGCTCTTATATGTTAGCTACGCTCCAGAGAA 448

RESULT 12
US-10-085-783A-26616
Sequence 26616, Application US/10085783A

GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Liew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2002
CURRENT APPLICATION NUMBER: US/10/085,783A
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2
SEQ ID NO 26616
LENGTH: 469
TYPE: DNA
ORGANISM: Human
US-10-085-783A-26616

Query Match 9.9%; Score 63.6; DB 6; Length 469;
Best Local Similarity 50.7%; Pred. No. 1.4e-06;
Matches 206; Conservative 0; Mismatches 194; Indels 6; Gaps 2;

QY 101 CCCAGAGCCCGGCTGCGCGCCAGCCGCGCCGCGCATGCTTCAGACCGGCTTTC 160
DB 35 CTCGGCGGCTCGCGAGCGCGGTTCCGTTCCCTCCCGCGCCGCGCATGAGTGGATGTC 94
QY 161 AAGCAGCGCGGAGCTTCCCGCGCGCTGTAAAGAGGTACAGAGATCCGCGAGCAGC 220
DB 95 AAGGAGAGCACTCCGCTGAGACAGAGTGGTGGTCCGAGAGATTCAGAGAAATAT 154
QY 221 CCCAGCAAAATCCCGGTATCATGAGCGCTCAAGAGGTGAGAGAGAGTGGTCCCTG 280
DB 155 CCCGACAGGCTTCCGCTGATGTGTGAAAGTGTCTAG--CTTCAGATTTTACATT 211
QY 281 GACAAGACCAAGTTTGTGTCGCCGAGCAGTGTCAACATGAGGAGTGGTCAAGTATC 340
DB 212 GACAAAGGAGTACTGTTGTTCTCATCTGATCTGATCTGATCTGATCTGATCTGATCT 271
QY 341 CGGCGCGGCTGAGCTGAGCCGCGAGCGAGCGCTTCTGCTGCTGAGACGAGCAGC 400
DB 272 AGGAAAGAGTCCAGCTTCTTCTGAAAGGAGATCTTCTTGTGTGATTAAGACAGTC 331
QY 401 ATGTGAGTGTGTCAGCGCCCATTCGCGAGCATCTACGAGAGAGAGAGAGAGAGG 460
DB 332 CCACAGTCCAGCTTACT--ATGGAGACGCTTTACGAGAGAGAGAGAGATGATGGA 388
QY 461 TTCTCTATATGTTCTACGCTTCCGAGAGAAACCTTGGCTTCTGAG 506
DB 389 TTCTATATATGTTCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 434

RESULT 13
US-10-085-783A-10191
Sequence 10191, Application US/10085783A
GENERAL INFORMATION:
APPLICANT: ChondroGene Inc.
APPLICANT: Liew, C.C.
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
FILE REFERENCE: 4231/2002
CURRENT APPLICATION NUMBER: US/10/085,783A
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/305,340
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/275,017
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 60/271,955
PRIOR FILING DATE: 2001-02-28
NUMBER OF SEQ ID NOS: 58994
SOFTWARE: PatentIn version 3.2
SEQ ID NO 10191

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2003, 07:57:30 : Search time 2040 seconds
(without alignments)
7624.934 Million cell updates/sec

Title: US-09-904-603-2
Perfect score: 640
Sequence: 1 CTCGCCGACGCCGCGCC.....GCCCCCTAGTCAGAGGCA 640

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlin:*
5: em_estlin:*
6: em_estlin:*
7: em_estlin:*
8: em_estlin:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estfun:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627.4	98.0	1149	12	BM925959 AGNCOURT
2	625.8	97.8	1074	12	BM919877 AGNCOURT
3	598.8	93.6	700	12	BI667032 603291206
4	593.4	92.7	705	10	BG715828 602675240

5	590.4	92.2	1094	12	BM919896
6	584.4	91.3	1115	12	BM924606
7	583.6	91.2	794	12	BI601673
8	582.4	91.0	738	12	BI824755
9	575.8	90.0	1151	12	BM924607
10	573.6	89.6	730	12	BI522709
11	563.4	88.0	676	12	BI918612
12	560	87.5	594	10	BE885329
13	534	83.4	1281	13	B0667721
14	514.4	80.4	560	14	CH111122
15	513	80.2	561	12	BI668829
16	501	78.3	766	10	BG708209
17	488	76.2	514	9	AI815709
18	485.2	75.8	528	10	BE712961
19	483.8	75.6	1009	13	BQ723424
20	482.6	75.4	614	14	CB450406
21	474.4	74.1	546	13	B0784127
22	474.4	74.1	546	13	BX115760
23	464	72.5	476	9	AW161293
24	447.8	70.0	623	14	CB286481
25	436.2	68.2	894	14	CB321417
26	433.4	67.7	633	10	AM962192
27	432.8	67.6	595	12	BM946584
28	428.4	66.9	513	13	BX115760
29	425.2	66.4	953	11	AK003716
30	424.4	66.3	655	13	BQ444944
31	422.4	66.0	961	13	B0590653
32	417.4	65.2	555	14	CB613461
33	417	65.2	501	10	BG383216
34	416.4	65.1	864	12	BI735232
35	415.2	64.9	688	14	CA338578
36	414.4	64.8	543	10	BE881349
37	413.2	64.6	963	11	AK015020
38	413.2	64.6	965	14	BY703975
39	409.4	64.0	680	14	CA322683
40	409.2	63.9	1065	10	BF300927
41	406.4	63.5	866	13	B0604726
42	406.4	63.5	954	11	AK003122
43	405.4	63.3	977	14	BY703479
44	402	62.8	418	10	BE712843
45	401.4	62.7	441	10	BE712850

ALIGNMENTS

RESULT 1
LOCUS BM925959 1149 bp mRNA linear EST 12-MAR-2002
DEFINITION AGNCOURT_6649827 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:576435
5' UTR mRNA sequence.
ACCESSION BM925959
VERSION BM925959.1 GI:19376434
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1149)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNC)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNC at:
<http://image.llnl.gov>
Plate: LLM12817 row: m column: 04
High quality sequence start: 3


```
Db      | 361 | CAGCATGCTGAGTGTGTCACGCCCATCGCGACATCTACGAGCAGGAGAAAGACGAGCA 420
Qy      | 457 | CGGCTCTCTATATAGTGTACGCTCCGAGAAACCTTGCGCTTCTAGCCAGCAGTAG 516
Db      | 421 | CGGCTCTCTATATAGTGTACGCTCCGAGAAACCTTGCGCTTCTAGCCAGCAGTAG 480
Qy      | 517 | GGGGGCTCGGCTCGGAGCTCGGGGGGCCCGGTCAGGCGCTTGCCAGAGAGCTTGTT 576
Db      | 481 | GGGGGCTCGGCTCGGAGCTCGGGGGGCCCGGTCAGGCGCTTGCCAGAGAGCTTGTT 540
Qy      | 577 | CCTGAAGTGAAGCTGCGCTTACCGTGTGTGGCTGGGCGAGGCATGTGCCCGCTAGTCAGAG 636
Db      | 541 | CCTGAAGTGAAGCTGCGCTTACCGTGTGTGGCTGGGCGAGGCATGTGCCCGCTAGTCAGAG 600
Qy      | 637 | GGCA 640
Db      | 601 | GGCA 604
```

```
RESULT 6
LOCUS   BM924606 1115 bp mRNA linear EST 12-MAR-2002
DEFINITION ACENCD00RT.6767666 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760990
5', mRNA sequence.
ACCESSION BM924606
VERSION   BM924606.1 GI:19374985
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1115)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs@mail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLM12808 row: c column: 07
           High quality sequence stop: 572.
           Location/Qualifiers
             1..1115
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5760990"
               /lab_host="DH10B"
               /note="organ: pooled colon, kidney, stomach; Vector:
               pCMV-Sport6; Site:1: North; Site:2: EcoRV (destroyed); RNA
               source anonymous pool of 3 colons, age 26 yo male, 49 yo
               female, 71 yo male colon; 46 yo male kidney, and pool of 2
               stomachs, 62 yo male and 70 yo female. Library is
               oligo-dT primed and directionally cloned (EcoRV site is
               destroyed upon cloning). Average insert size 1.4 kb,
               insert size range 1-3 kb. Library is normalized and
               enriched for full-length clones and was constructed by C.
               Gruber (Invitrogen). Research Genetics tracking code
               023. Note: this is a NIH_MGC Library."
```

```
BASE COUNT 125 a 375 c 389 g 114 t 112 others
ORIGIN
```

```
Query Match 91.3% Score 584.4 DB 12: Length 1115:
Best Local Similarity 99.5% Pred. No. 52e-111:
Matches 597: Conservative 0: Mismatches 1: Indels 2: Gaps 1:
```

```
Qy      | 43 | CCTTGAGCGGAGAGCGGGAAGGCC--GGAGGCCGCAAGCGAGACATACGCGCG 109
Db      | 1 | CCTTGAGCGGAGAGCGGGAAGGCCCGCGAGGCCGCAAGCGAGACATACGCGCG 60
Qy      | 101 | CCGAGAGGCCCGGCGCTGGGCGCGGAGCGCGCGCGGCGGCGGCGGCGGCGGCGG 160
Db      | 61 | CCGAGAGGCCCGGCGCTGGGCGCGGAGCGCGCGCGGCGGCGGCGGCGGCGGCGG 120
Qy      | 161 | AAGAGAGCGGCGGAGAGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 220
Db      | 121 | AAGAGAGCGGCGGAGAGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Qy      | 221 | CCGAGCAAAATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 280
Db      | 181 | CCGAGCAAAATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Qy      | 281 | GACAGACCCAGAGTTTGGTGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 440
Db      | 241 | GACAGACCCAGAGTTTGGTGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 400
Qy      | 341 | CGGCGCGCGCGCTGCGAGCTGAGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 400
Db      | 301 | CGGCGCGCGCGCTGCGAGCTGAGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 460
Qy      | 401 | ATGCTGAGTGTGTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 460
Db      | 361 | ATGCTGAGTGTGTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Qy      | 461 | TTCTCTATATGCTTACGCGCTGCGAGGAAACCTTGGGCTTGTGAGCGGAGGAG 520
Db      | 421 | TTCTCTATATGCTTACGCGCTGCGAGGAAACCTTGGGCTTGTGAGCGGAGGAG 480
Qy      | 521 | GCTGCGCTGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 580
Db      | 481 | GCTGCGCTGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Qy      | 581 | AACTGAGCTGCGCTTACCGCTGTGGGCTGGGCGGAGGCGGCGGCGGCGGCGGCGG 640
Db      | 541 | AACTGAGCTGCGCTTACCGCTGTGGGCTGGGCGGAGGCGGCGGCGGCGGCGGCGG 600
```

```
RESULT 7
LOCUS   B1601673 794 bp mRNA linear EST 07-SEP-2001
DEFINITION 603249912P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301554 5',
mRNA sequence.
ACCESSION B1601673
VERSION   B1601673.1 GI:15494612
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 794)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs@mail.nih.gov
           Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
           cDNA Library Preparation: Michael J. Brownstein (NIH/NIH).
           Toshiyuki and Piero Carninci (RIKEN)
           National Institutes of Health, Mammalian Gene Collection (MGC)
           DNA Sequencing by: Invitrogen Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLM11763 row: h column: 93
           High quality sequence stop: 751.
           Location/Qualifiers
             1..794
               /organism="Homo sapiens"
               /mol_type="mRNA"
```

```
BASE COUNT 125 a 375 c 389 g 114 t 112 others
ORIGIN
```

```
Query Match 91.3% Score 584.4 DB 12: Length 1115:
Best Local Similarity 99.5% Pred. No. 52e-111:
Matches 597: Conservative 0: Mismatches 1: Indels 2: Gaps 1:
```

[illegible]

KEYWORDS	EST	Homo sapiens (human)
SOURCE		Homo sapiens
ORGANISM		Homo sapiens (human)
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS		1 (bases 1 to 738)
TITLE		NIH-MGC http://mgi.nci.nih.gov/
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT		Unpublished
		Contact: Robert Strausberg, Ph.D.
		Email: cga@b-remail.nih.gov
		Tissue Procurement: Life Technologies, Inc.
		cDNA Library Preparation: Life Technologies, Inc.
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
		DNA Sequencing by: Incyte Genomics, Inc.
		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at:
		http://image.lnl.gov
		Plate: L14M11435 row: 9 column: 10
		High quality sequence stop: 718.
FEATURES		location/qualifiers
source		1..738
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="IMAGE:5174817"
		/lab_host="DH10B"
		/clone.lib="NIH-MGC_115"
		/note="Organ: pooled brain, lung, testis; Vector: PCMV-SpOrie; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."
BASE COUNT		132 a 244 c 236 g 126 t
ORIGIN		
Query Match		91.0%; Score 582.4; DB 12; Length 738;
Best Local Similarity		99.7%; Pred. No. 1.3e-110;
Matches 594; Conservative		0; Mismatches 1; Indels 1; Gaps 1;
QY	46	TTGAGCGGAGGCGCGGAG-CCCGGAGACCCCAACCGCAGACATCCCGGCGCCCA 104
Db	1	TTGAGCGGAGGCGCGGAGCGCCCGGAGACCCCAACCGCAGACATCCCGGCGCCCA 60
QY	105	GAGCCCGGCGCTCGGCGCCCAAGCGGGCGCGCGAGTCCTCAGACCGGCGCTTTCAAGC 164
Db	61	GAGCCCGGCGCTCGGCGCCCAAGCGGGCGCGCGAGTCCTCAGACCGGCGCTTTCAAGC 120
QY	165	AGCGGCGAGCTTGCGCGAGCGGTGAAGGATACAGCAGATCCGAGCAGACCCCA 224
Db	121	AGCGGCGAGCTTGCGCGAGCGGTGAAGGATACAGCAGATCCGAGCAGACCCCA 180
QY	225	GCAAAATCCGCGTATCATGAGCGCTTCAAGAGGTGAGAAACAGCTGCGCTCTGGACA 284
Db	181	GCAAAATCCGCGTATCATGAGCGCTTCAAGAGGTGAGAAACAGCTGCGCTCTGGACA 240
QY	285	AGACCAAGTTTGGTCCCGGACCATATGCAACATGAGCGAGTTGGTCAAGATCATCGGC 344
Db	241	AGACCAAGTTTGGTCCCGGACCATATGCAACATGAGCGAGTTGGTCAAGATCATCGGC 300
QY	345	GCGCGCTGACAGCTAAACCCCAAGAGGCGCTTCTCGTGTGTAACAGACAGACATGG 404
Db	301	GCGCGCTGACAGCTAAACCCCAAGAGGCGCTTCTCTCTGCTGTTAACCAGCAGCATGG 360
QY	405	TGAGTGTGTCCACGCCCATCGCGACATCTTACGAGCAGAGAAAGAGAGGAGCGCTTCC 464
Db	361	TGAGTGTGTCCACGCCCATCGCGACATCTTACGAGCAGAGAAAGAGAGGAGCGCTTCC 420

QY 465 TCATATAGTCTACGCTCCCGAGAAACCTTCGGCTTCTGAGCCAGAGTAGAGGGGGCTC 524
 |||||||
 Db 421 TCATATAGTCTACGCTCCCGAGAAACCTTCGGCTTCTGAGCCAGAGTAGAGGGGGCTC 480
 QY 525 GGCTTGGAGTGGGGGGCCCGGTGACAGCCCTCCCGAGAGAGCTTGTGTTCTGAAC 584
 |||||||
 Db 481 GGCTTGGAGTGGGGGGCCCGGTGACAGCCCTCCCGAGAGAGCTTGTGTTCTGAAC 540
 QY 585 GAGTGTCTCTACGCTGGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGGCA 640
 |||||||
 Db 541 GAGTGTCTCTACGCTGGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGGCA 596

 RESULT 9
 HM924607 mRNA linear EST 12-MAR-2002
 LOCUS HM924607
 DEFINITION 5', mRNA sequence.
 ACCESSION BM924607
 VERSION BM924607.1 GI:15374986
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1151)
 NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNI)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNI at:
 http://image.llnl.gov
 Plate: LLMI12808 row: 0 column: 08
 High quality sequence start: 44
 High quality sequence stop: 718.
 Location/Qualifiers
 1..1151
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5760991"
 /lab_host="DH10B"
 /clone_lib="NIH-MGC_116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-Sport6; Site1: NotI; Site2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 25 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH-MGC Library."
 BASE COUNT 190 a 317 c 409 g 196 t 39 others
 ORIGIN
 Query Match 90.0%; Score 575.8; DB 12; Length 1151;
 Best Local Similarity 99.8%; Pred. No. 3,le-109;
 Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 65 CCGCGGAGCCCGCAACCGGAGACATCCCGGCGCCGAGAGCCCGGCTGCGCGCC 124
 |||||||
 Db 54 CCGCGGAGCCCGCAACCGGAGACATCCCGGCGCCGAGAGCCCGGCTGCGCGCC 113
 QY 125 AGCGGCGCGCGCGGATGCGCTCAGACCGGCGCTTCAAGCAGCGGCGTACGTTGCGCGAC 184
 |||||||

Db 114 AGCGGCGCGCGCGGATGCGCTCAGACCGGCGCTTCAAGCAGCGGCGGAGCTTCGCGAC 174
 QY 185 CGCTGTAGGAGGATACAGTACATCGGTGAGTACAGACCCGACGAAATTCGGGTATCAT 244
 |||||||
 Db 174 CGCTGTAGGAGGATACAGTACATCGGTGAGTACAGACCCGACGAAATTCGGGTATCAT 233
 QY 245 GAGGCGTACAGAGGATGAGACAGCATGTCGCGTCTGGAGCAGAGCAAGCTTTTCTGCTG 304
 |||||||
 Db 234 GAGGCGTACAGAGGATGAGACAGCATGTCGCGTCTGGAGCAGAGCAAGCTTTTCTGCTG 293
 QY 305 GAGCATGTCAACATGAGTACGATGTTGATCAATATCCGCGCTGCTGATGATCAATGCT 364
 |||||||
 Db 294 GAGCATGTCAACATGAGTACGATGTTGATCAATATCCGCGCTGCTGATGATCAATGCT 353
 QY 365 AGCGAGGCTTCTTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 424
 |||||||
 Db 354 AGCGAGGCTTCTTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 413
 QY 425 GCGGACATCTACGAGCAG 484
 |||||||
 Db 414 GCGGACATCTACGAGCAG 473
 QY 485 CAGGAAATCTTCGGCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
 |||||||
 Db 474 CAGGAAATCTTCGGCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 533
 QY 545 CCGGTACAGGCTTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 604
 |||||||
 Db 534 CCGGTACAGGCTTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 593
 QY 605 GCGTGGGAGGAGCATGTGCGGCTGAGTACAGAGGCA 640
 |||||||
 Db 594 GCGTGGGAGGAGCATGTGCGGCTGAGTACAGAGGCA 629

 RESULT 10
 B1522709 mRNA linear EST 29-AUG-2001
 LOCUS B1522709
 DEFINITION 603175758P1 NIH-MGC_121 Homo sapiens cDNA clone IMAGE:5240007 5',
 mRNA sequence.
 ACCESSION B1522709
 VERSION B1522709.1 GI:15347501
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 730)
 NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNI)
 DNA Sequencing by: Lucite Synomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNI at:
 http://image.llnl.gov
 Plate: LLMI1605 row: 0 column: 16
 High quality sequence stop: 740.
 Location/Qualifiers
 1..730
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5240007"
 /lab_host="DH10B"
 /clone_lib="NIH-MGC_121"
 /note="Organ: brain; Vector: pCMV-Sport6; Site1: NotI;
 Site2: EcoRV (destroyed); RNA source anonymous pool of 4
 fetal brains, female and 20 weeks, female and 24 weeks,

and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."

BASE COUNT 129 a 245 c 242 g 114 t
ORIGIN

Query Match 89.6%; Score 573.6; DB 12; Length 730;
Best Local Similarity 98.0%; Pred. No. 8.3e-109;
Matches 623; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

9 GCCGAGCCGCGCTCTAGCGCGAGCCCGGAGCCCTTGGAGCGGAGCGGAGCCCG 68
1 GCCGAGCCGCGCTCTAGCGCGAGCCCGGAGCCCTTGGAGCGGAGCGGAGCCCG 60
69 G-GAGCCCGCGAGCGAGCATCCCGCGCGCGAGAGCC-CGGGCTGGCGCGCCAG 126
61 GCGGAGCCCGCGAGCGAGCATCCCGCGCGCGAGAGCGCGCGCGCGCGCGCCAG 120
127 CCGGCGCGCGCGAGCGAGCATCCCGCGCGCGAGAGCGCGAGCGCGCGCGCGCG 186
121 CCGGCGCGCGCGAGCGAGCATCCCGCGCGCGAGAGCGCGAGCGCGCGCGCGCG 180
187 CTGTAGAGAGTACAGAGATCCGCGAGCGAGCGCGAGCGCGAGCGCGAGCGCG 246
181 CTGTAGAGAGTACAGAGATCCGCGAGCGAGCGCGAGCGCGAGCGCGAGCGCG 240
247 GCGGTACAGAGTACAGAGATCCCGCGCGCGAGCGCGAGCGCGAGCGCGAGCG 306
241 GCGGTACAGAGTACAGAGATCCCGCGCGCGAGCGCGAGCGCGAGCGCGAGCG 300
307 CCATGTCAACATGAGCGAGTGTGTCAGATCATCCGCGCGCGCGCGCGCGCGCG 365
301 CCATGTCAACATGAGCGAGTGTGTCAGATCATCCGCGCGCGCGCGCGCGCGCG 360
366 CCGAGGCGCTTCTCTCTGCTGTAACAGCAGCAGCATGAGTGTGTCAGCGCGCG 425
361 CCGAGGCGCTTCTCTCTGCTGTAACAGCAGCAGCATGAGTGTGTCAGCGCGCG 420
426 CCGAGGCGCTTCTCTCTGCTGTAACAGCAGCAGCATGAGTGTGTCAGCGCGCG 484
421 CCGAGGCGCTTCTCTCTGCTGTAACAGCAGCAGCATGAGTGTGTCAGCGCGCG 480
485 CAGGAACCTTCTCTCTGCTGTAACAGCAGCAGCATGAGTGTGTCAGCGCGCG 544
481 CAGGAACCTTCTCTCTGCTGTAACAGCAGCAGCATGAGTGTGTCAGCGCGCG 540
545 CCGGCGCGCGCGAGCGAGCATCCCGCGCGCGAGAGCGCGAGCGCGCGCGCGCG 604
541 CCGGCGCGCGCGAGCGAGCATCCCGCGCGCGAGAGCGCGAGCGCGCGCGCGCG 600
605 GCGTGGCGAGGATGTGCGCGCGCGAGAGCGCGAGCGCGAGCGCGAGCGCGCG 640
601 GCGTGGCGAGGATGTGCGCGCGCGAGAGCGCGAGCGCGAGCGCGAGCGCGCG 636

RESULT 11
LOCUS B1918612 676 bp mRNA linear EST 16-OCT-2001
DEFINITION 60317655F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241013 5',
ACCESSION B1918612
VERSION B1918612
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 676)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL1607 row: m column: 14
High quality sequence start: 2
High quality sequence stop: 671.
Location/Qualifiers
1. 676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5241013"
/lab_host="DH10B"
/clone.lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."

FEATURES
source

BASE COUNT 123 a 223 c 217 g 113 t
ORIGIN

Query Match 88.0%; Score 563.4; DB 12; Length 676;
Best Local Similarity 98.8%; Pred. No. 1.1e-106;
Matches 578; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

56 GCGCGGAGCGCGCGAGCGCGCGAGCGCGAGCGCGCGCGCGCGCGCGCGCG 115
3 GATGTGAGCGCGCGAGCGCGCGAGCGCGAGCGCGCGCGCGCGCGCGCGCG 62
116 TCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 175
63 TCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122
176 TTGCG 235
123 TTGCG 182
236 GTGATCATGAGCGCGTACAGAGGTTGTCAGATCATCCGCGCGCGCGCGCG 295
183 GTGATCATGAGCGCGTACAGAGGTTGTCAGATCATCCGCGCGCGCGCGCG 242
296 TTGCTCCCGGAGCATGTACATGAGCGAGTTGTCAGATCATCCGCGCGCGCG 355
243 TTGCTCCCGGAGCATGTACATGAGCGAGTTGTCAGATCATCCGCGCGCGCG 302
356 CTGAACCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
303 CTGAACCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 362
416 ACGCCCATCGCGCGCATGTACAGAGGTTGTCAGATCATCCGCGCGCGCGCG 475
363 ACGCCCATCGCGCGCATGTACAGAGGTTGTCAGATCATCCGCGCGCGCGCG 422
476 TACGCTTCCAGGAACCTTTCGCTTGTGAGCCAGAGTGTAGGGGCTCGGCTGG 535
423 TACGCTTCCAGGAACCTTTCGCTTGTGAGCCAGAGTA-GGGGCTCGGCTGG 481
536 CCGGCG 595

OY	1	CTCCCGCAGCCGACCGCGCTGTCAAGCGCAGCGCCCGGAGCGCTTTGAGCGCGAGAGCG	60
Db	43	CTCCCGCAGCCGACCGCGCTGTCAAGCGCAGCGCCCGGAGCGCTTTGAGCGCGAGAGCG	102
OY	61	GGAG-CGCCGGAGCCCGCCAAACCGCAGACACATCCCGCGCCCCAGAGCGCCGCGCTGG	119
Db	103	GGAGCGCCCGGAGCGCCCGCCAAACCGCAGACACATCCCGCGCCCCAGAGCGCCGCGCTGG	162
OY	120	CGCCGAGCGGGCCCGCGCGATGCCCTAGACCGGCTTTCAAGCAGCGCGGAGGCTTGG	179
Db	163	CGCCGAGCGGGCCCGCGCGATGCCCTAGACCGGCTTTCAAGCAGCGCGGAGGCTTGG	222
OY	180	CCGACCGCTGTAAAGGAGGTACAGCAGATCCGGGACAGACGCCACCAAAATCCCGGTGA	239
Db	223	CCGACCGCTGTAAAGGAGGTACAGCAGATCCGGGACAGACGCCACCAAAATCCCGGTGA	282
OY	240	TCATCGACGCGTACAAAGGTGAGAGCAGCTCCCGCTCTGGACAAAGCAAGATTTTTGG	299
Db	283	TCATCGACGCGTACAAAGGTGAGAGCAGCTCCCGCTCTGGACAAAGCAAGATTTTTGG	342
OY	300	TCCCGGACCATGTCAACATGAGCGAGTTGGTCAAGATCTCCGGCGCGCGCTTGACGTGA	359
Db	343	TCCCGGACCATGTCAACATGAGCGAGTTGGTCAAGATCTCCGGCGCGCGCTTGACGTGA	402
OY	360	ACCCGAGCGAGCGCTTCTCCGCTGGGTGAACACACAGCATGTGGATGTGTCCAGCC	419
Db	403	ACCCGAGCGAGCGCTTCTCCGCTGGGTGAACACACAGCATGTGGATGTGTCCAGCC	462
OY	420	CCATCGCGAGCATCTACAGCAGAGAGAAAGACGAGCGGCTTCCCTATATGGTGTAG	479
Db	463	CCATCGCGAGCATCTACAGCAGAGAGAGAAAGACGAGCGGCTTCCCTATATGGTGTAG	522
OY	480	CCTCCGAGAAACCTTGGGCTTCTGAGCCAGCAGTAGGCGGGGCTCGGCTTGGAGTCGG	539
Db	523	CCTCCGAGAAACCTTGGGCTTCTGAGCCAGCAGTAGGCGGGGCTCGGCTTGGAGTCGG	582
OY	540	GGGCGCGGCTCAGG	553
Db	583	GCGACCCGAGCTGG	596

RESULT	14
CBI11122	
LOCUS	CBI11122
DEFINITION	CBI11122 560 bp mRNA linear EST 28-JAN-2003
ACCESSION	K-E010152658 LSHLK1 Homo sapiens cDNA clone LSHLK1-8-G02 5' mRNA sequence.
VERSION	CBI11122
KEYWORDS	CBI11122.1 GI:27936929
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 560)
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Kim,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE	Zic Frontier Korean EST Project 2001
JOURNAL	Unpublished
COMMENT	Contact: Kim YS

```

Genome Research Center
Korea Research Institute of Biosci.
52 Eosun-dong Yuseong-gu, Daejeon
Tel : +82-42-860-4470
Fax : +82-42-860-4409
Email: yongsung@gmail.kr,krb,rc.kr
Plate: 8 row: 5 column: 02
High quality sequence stop: 560.
FEATURES
    source            location/Qualifiers
    1..560
    /organism="Homo sapiens"
    /mol_type="mRNA"

```

```

/adb_xref="taxon:3606"
/clone="L5HLK1-8-G02"
/sex="M"
/cell_line="HLK-1"
/lab_host="Top10f"
/clone_lib="L5HLK1"
/note="Organ: Liver; Vector: PCNS-D2; Site:1: EcoRI;
Site:2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deapped
with tabacco acid pyrophosphatase (TAP). The deapped
Intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10f by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

```

Query Match	80.48;	Score 514.4;	DB 14;	Length 560;
Best Local Similarity	99.68;	Pred. No. 1.4e-96;		
Matches 526;	Conservative	0;	Mismatches 1;	Indels 1;
			Gaps	1;

QY	1	CTCCCGAGAGCCGCGAGCCGCGCTGGTCTCAGACGGAGAGCGGAGACCTTTAGCGGCGAGGGCG	60
Db	33	CTCCCGAGAGCCGCGAGCCGCGCTGGTCTCAGCGCAAGGCCCGGAGGCCCTTGAAGCGCGAGGGCG	92
QY	61	GGAG-CCCCGGAGGCCCCCAAAACCCGACACATCCCGCGGCCCCGAGAGCCCGGCTCGG	119
Db	93	GGAGCCCCCGGAGGCCCCCAAAACCCGACACATCCCGCGGCCCCGAGAGCCCGGCTCGG	152
QY	120	CGCCAGACCGGGCCCGCGCGATGGCTTCAGACCGGCGCTTTCAGGAGCGGGGAGACTTCG	179
Db	153	CGCCAGACCGGGGCCCGCGCGATGGCTTCAGACCGGCGCTTTCAGGAGCGGGGAGACTTCG	212
QY	180	CCGACCGCTGTAAAGGAGGTACAGAGATCCGCGACACGACCCGACGAAATCCCGGTGA	239
Db	213	CCGACCGCTGTAAAGGAGGTACAGAGATCCGCGACACGACCCGACGAAATCCCGGTGA	272
QY	240	TCATCGAGCGCTACAAGGGTGAGAAAGCAGTGCCTCGCTGGACAAGACCAAGTTTTTGG	299
Db	273	TCATCGAGCGGTACAAAGGTTGAGAAAGCAGTGCCTCGCTGGACAAGACCAAGTTTTTGG	332
QY	300	TCCCGGACCATGTAAACATGAGCGAGTTGGTTCAGAGATATCCGGCGCCGCTGCAGCTGA	359
Db	333	TCCCGGACCATGTTCMACATGAGCGAGTTGGTTCAGAGATATCCGGCGCCGCTGCAGCTGA	392
QY	360	ACCCAGAGCAGGCCCTTCTCCTCGTGGGGAACGACGACAGATGGTGATGTGTCACAGC	419
Db	393	ACCCAGAGCAGGCCCTTCTCCTCGTGGTGAACCGACACAGATGGTGATGTGTCACAGC	452
QY	420	CCATCGCGGACATCTACGACAGGAGAGAAGACGAGAGCGGCTTCCTATATGGTCTACG	479
Db	453	CCATCGCGGACATCTACGACAGGAGAGAAGACGAGAGCGGCTTCCTATATGGTCTACG	512
QY	480	CCATCCAGGAACCTTGGGCTTCTGAGCCACGACACTAAGGGGGGCTCGGC	527
Db	513	CCATCCAGGAACCTTGGGCTTCTGAGCCACGACACTAAGGGGGGCTCGGC	560

LOCUS	DEFINITION	LOCUS	DEFINITION
RESULT 15 B1668829	561 bp mRNA linear EST 12-SEP-2001 6033294786F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5313956 5', mRNA sequence.	B1668829	B1668829
ACCESSION	B1668829	ACCESSION	B1668829
VERSION	B1668829.1 GI:15583062	VERSION	B1668829.1 GI:15583062

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 561)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ega@bts-remail.nih.gov
Tissue Procurement: Miklos Palakovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11795 row: 1 column: 21
High quality sequence stop: 558.

FEATURES
Source
Location/Qualifiers
1..561
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5313956"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_id="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 109 a 200 c 173 g 79 t
ORIGIN

Query Match 80.2%; Score 513; DB 12; Length 561;
Best Local Similarity 99.5%; Pred. No. 2.7e-96;
Matches 546; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 CTCCTCCAGCCGACACCGCGTGTCTCAGCGCGAGCCCGGAGCCCTTGAGCGCGAGCGC 60
DB 13 CTCCTCCAGCCGACACCGCGTGTCTCAGCGCGAGCCCGGAGCCCTTGAGCGCGAGCGC 72
QY 61 GGAG-CCCGGAGCGCCCAACCGGACACATCCCGCGCCCGCAGAGCCCGCGCTGCG 119
DB 73 GGAGCCCGCGAGCCCGCAACCGGACACATCCCGCGCCCGCAGAGCCCGCGCTGCG 132
QY 120 CGCCAGCGCGCCCGC-GCGATGCGCTCAGACCGCGCTTCAAGCAGCGCGAGCTTC 178
DB 133 CGCCAGCGCGCGCGCTGCGATGCGCTCAACCGCGCTTCAAGCAGCGCGAGCTTC 192
QY 179 GCGCAGCGCTGTAAAGAGGTACAGACATCCGCGACACGACCCGAGCAAAATCCCGGTG 238
DB 193 GCGCAGCGCTGTAAAGAGGTACAGACATCCGCGACACGACCCGAGCAAAATCCCGGTG 252
QY 239 ATCATCGAGCGCTCAAGAGGTGAGAGAGTCCCGCTCGAGCAAGACCAAGTTTTT 298
DB 253 ATCATCGAGCGCTCAAGAGGTGAGAGAGTCCCGCTCGAGCAAGACCAAGTTTTT 312
QY 299 GTCCCGGACATGTCAACATGAGGAGTGTGTCAAGATCATCCGCGCGCTGAGCTG 358
DB 313 GTCCCGGACATGTCAACATGAGGAGTGTGTCAAGATCATCCGCGCGCTGAGCTG 372
QY 359 AACCCAGCAGCGCTTCTCTCTGTGTGTAACAGCAGACAGCATGGTGTGTGTCAAG 418
DB 373 AACCCAGCAGCGCTTCTCTCTGTGTGTAACAGCAGACAGCATGGTGTGTGTCAAG 432

QY 419 CCATCCGCGACATCTAGAGTACGAGAAAGACGAGAC-GGCTTCTCTATATATGTTCA 477
DB 433 CCATCCGCGAGCATCTAGAGTACGAGAAAGACGAGAGCGGCTTCTCTATATATGTTCA 492
QY 478 CCCTCCGAGAAACCTTCGCTCTCTGAGCCAGCAGTACGCGGCTCGGCTCTGAGTGG 537
DB 493 CCCTCCGAGAAACCTTCGCTCTCTGAGCCAGCAGTACGCGGCTCTGAGTGGTGGTGG 552
QY 538 GGGGGCCCC 546
DB 553 GGGGGCCCC 561

Search completed: August 15, 2003, 09:30:40
Job time : 2046 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2003, 08:56:30 : Search time 249 Seconds
(without alignments)
5752.136 Million cell updates/sec

Title: US-09-904-603-2
Perfect score: 640
Sequence: 1 CTCGCCGAGCGCAGCGGCC.....GCCCCCTACTGACAGGCA 640

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 118970152 residues
Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query	Match Length	ID	Description
1	640	100.0	640	11	US-09-904-603-2
2	255.2	39.9	570	9	US-09-864-761-12594
3	238.6	37.3	861	10	US-09-917-800A-1714
4	238.6	37.3	861	12	US-10-205-194-72
5	218.8	34.2	497	11	US-09-918-995-1874
6	215	33.6	768	14	US-10-037-270-707
7	215	33.6	2250	9	US-09-925-301-209
8	214.4	33.5	2199	14	US-10-277-156-3
9	192.4	30.1	537	10	US-09-917-800A-1201
10	163	25.5	163	9	US-09-864-761-28415
11	108.4	16.9	201	14	US-10-277-156-1
12	95.6	14.9	474	9	US-09-864-761-885
13	71.4	11.2	155	9	US-09-864-761-17665
14	71.4	11.2	975	10	US-09-917-800A-1711
15	71.4	11.2	975	14	US-10-205-342-24
16	58.8	9.2	500	11	US-09-918-995-1734

17	58.8	9.2	726	13	US-10-027-632-148570	Sequence 148570, App
18	57.2	8.9	363	10	US-09-938-842A-461	Sequence 461, App
19	54.4	8.5	634	9	US-09-770-119-663	Sequence 663, App
20	50.8	7.9	15500	9	US-09-764-869-2109	Sequence 2109, App
21	50.8	7.9	15500	14	US-10-091-504-2109	Sequence 2109, App
22	50.4	7.9	3828	14	US-10-156-761-2959	Sequence 2959, App
23	50.4	7.9	9025608	12	US-10-156-761-1	Sequence 1, App
24	49.6	7.8	487	14	US-10-205-194-161	Sequence 161, App
25	49.2	7.7	759	10	US-09-919-580-825	Sequence 825, App
26	49	7.7	2038	14	US-10-037-270-597	Sequence 597, App
27	47.2	7.4	458	11	US-09-918-995-28426	Sequence 28426, App
28	47.2	7.4	483	10	US-09-919-580-198	Sequence 198, App
29	47.2	7.4	498	14	US-10-171-581-105	Sequence 105, App
30	47.2	7.4	519	10	US-09-736-457-78	Sequence 78, App
31	47.2	7.4	519	10	US-09-902-941-78	Sequence 78, App
32	47.2	7.4	519	10	US-09-849-626-78	Sequence 78, App
33	47.2	7.4	519	11	US-09-476-300-78	Sequence 78, App
34	47.2	7.4	519	14	US-10-017-754-78	Sequence 78, App
35	47	7.3	891	14	US-10-156-761-4852	Sequence 4852, App
36	46.6	7.3	975	13	US-10-027-632-9165	Sequence 9165, App
37	46.6	7.3	975	13	US-10-027-632-9165	Sequence 9165, App
38	46.6	7.3	975	13	US-10-027-632-9167	Sequence 9167, App
39	46.2	7.2	1179	14	US-10-156-761-6762	Sequence 6762, App
40	46	7.2	571	13	US-10-027-632-27666	Sequence 27666, App
41	45.8	7.2	929	9	US-09-814-122-16	Sequence 16, App
42	45.6	7.1	484	11	US-09-918-995-6964	Sequence 6964, App
43	45.2	7.1	450	10	US-09-910-664-59	Sequence 59, App
44	45.2	7.1	1155	14	US-10-156-761-5751	Sequence 5751, App
45	45.2	7.1	9025608	14	US-10-156-761-1	Sequence 1, App

ALIGNMENTS

RESULT 1
US-09-904-603-2
Sequence 2, Appl
Publication No. US2004009461ZAI
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
GOLLY, Surya K.
TITLE OF INVENTION: N-VEL MICRODOUBLE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/904,603
FILING DATE: 12-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 96,749
REFERENCE/DOCKET NUMBER: PR-0211 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 640 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: JYHNOT03
CLONE: 1441378
SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-904-603-2

Query Match 100.0%; Score 640; DB 11; Length 640;
Best Local Similarity 100.0%; Pred. No. 7.6e-157;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTCCCGAGCCGACCGCCGCTGCTCAGCGGACCCCGAGACCTTGAGCGGAGGCGC 60
DB 1 CTCCCGAGCCGACCGCCGCTGCTCAGCGGAGCCCGAGACCTTGAGCGGAGGCGC 60
QY 61 GGAGCCCCGGAGCCCCCAACCGCAGACATCCCGCGCCCGCAGAGCCCGGCTGGCC 120
DB 61 GGAGCCCCGGAGCCCCCAACCGCAGACATCCCGCGCCCGCAGAGCCCGGCTGGCC 120
QY 121 GCCAGCGGGGCGCGCGATGCCCTCAGACCGGCGCTTCAAGAGGCGCGAGCTTCGC 180
DB 121 GCCAGCGGGGCGCGCGATGCCCTCAGACCGGCGCTTCAAGAGGCGCGAGCTTCGC 180
QY 181 CGACCGCTGTAAAGAGGTACAGACATCCGCGACCGACCCCGCAAAATCCCGGTGAT 240
DB 181 CGACCGCTGTAAAGAGGTACAGACATCCGCGACCGACCCCGCAAAATCCCGGTGAT 240
QY 241 CATGAGCGGTACAAAGGTGAGAGCAGCTCCGCTCTGACAAAGCAAGTTTGTGT 300
DB 241 CATGAGCGGTACAAAGGTGAGAGCAGCTCCGCTCTGACAAAGCAAGTTTGTGT 300
QY 301 CCCGACCATGTCAACATGAGCGAGTTGTCAGATCATCCGCGCGCGCTGACGTGAA 360
DB 301 CCCGACCATGTCAACATGAGCGAGTTGTCAGATCATCCGCGCGCGCTGACGTGAA 360
QY 361 CCCGACCATGTCAACATGAGCGAGTTGTCAGATCATCCGCGCGCGCTGACGTGAA 360
DB 361 CCCGACCATGTCAACATGAGCGAGTTGTCAGATCATCCGCGCGCGCTGACGTGAA 360
QY 421 CATGCGGACATGTACAGAGGAGGAAAGACGAGCGCTTCTCTATATAGCTTACGC 480
DB 421 CATGCGGACATGTACAGAGGAGGAAAGACGAGCGCTTCTCTATATAGCTTACGC 480
QY 481 CTCCAGAGAAACCTTCGCTCTGAGCAGAGTAGGGGCTGGCTGGAGATCGGGG 540
DB 481 CTCCAGAGAAACCTTCGCTCTGAGCAGAGTAGGGGCTGGCTGGAGATCGGGG 540
QY 541 GGCCCGGCTAGAGCCCTGCCCCAGAGAGCTTCTGTTCTGAATGAGCTGCTTACCGT 600
DB 541 GGCCCGGCTAGAGCCCTGCCCCAGAGAGCTTCTGTTCTGAATGAGCTGCTTACCGT 600
QY 601 GGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGA 640
DB 601 GGTGGGCTGGGAGGAGCATGTGCCCCCTAGTCAGAGGGA 640
```

RESULT 2

US-09-864-761-12594
Sequence 12594 Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR APPLICATION NUMBER: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312

```
QY 340 CCGGCGCGCGCTGAGCTGAACCCAGCAGGCTTCTCTGCTGTGAACGACAG 399
DB 311 CAGCGCGCGCTGAGCTGAACCCAGCAGGCTTCTCTGCTGTGAACGACAG 370
QY 400 CATGCTAGTGTCTCCAGCGCCCATCGGGGACATCTAGAGAGGAGGAAAGCGAGCGG 459
DB 371 CATGCTAGTGTCTCCAGCGCCCATCGGGGACATCTAGAGAGGAGGAAAGCGAGCGG 430
QY 460 CTTCCTATATAGTCTACAGCTCCAGAAACCTTGAGGCTTCTGAGCCAGACATAGGGG 519
DB 431 CTTCCTATATAGTCTACAGCTCCAGAAACCTTGAGGCTTCTGAGCCAGACATAGGGG 490
QY 520 GGTCTGCTGAGAGTGGGGGGCGCGGTGAGGCTTCTGAGCCAGAGAGCTTGTGCT 579
DB 491 GGTCTGCTGAGAGTGGGGGGCGCGGTGAGGCTTCTGAGCCAGAGAGCTTGTGCT 550
QY 580 GAATGAGCTGCTTACCG 599
DB 551 GAATGAGCTGCTTACCG 570
```

Query Match 39.9%; Score 255.2; DB 9; Length 570;
Best Local Similarity 98.8%; Pred. No. 5.7e-57;
Matches 257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

US-09-864-761-12594
FEATURE: Homo sapiens
OTHER INFORMATION: MAP TO AL118520.12
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.67
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 12594
LENGTH: 570
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: Homo sapiens
OTHER INFORMATION: MAP TO AL118520.12
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.67
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
US-09-864-761-12594


```

CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1874
LENGTH: 497
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(497)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1874

```

```

Query Match          34.2%; Score 218.8; DB 11; Length 497;
Best Local Similarity 66.8%; Pred. No. 1.6e-47;
Matches 310; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

```

QY 35 CCCCAGAGCCCTTGAGCGCGAGCGCGAGCCCGGAGCCCGCAAAACCGCAGACATCC 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 CCACTGAGACCTGATGATGCGGTAAATGCGAGCGCGCGCGCGCGCGCGCGCGCG 72
QY 95 CCGCGCCCGCAGAGCCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GGAACCTCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 132
QY 155 CTTTGAAGAGGGGGGAGCTTCCGCGACCGCTGTAAGAGGTACAGACATCCGGAC 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 ACCCTTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 192
QY 215 CACGACCCCGCAGAAATCCCGGTATCATGAGCGCTACAGAGGTGAGAGCGAGCT 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 CACATCCCAACCAAAATCCCGGTATCATGAGAGGTGAGAGCGAGAGCTTCT 252
QY 275 GTCTGAGCAAGCAAGTTTGTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GTTCTGATTAACCAAGTTCTTGTACCTGACCATGTCATGATGAGTGCATCAAG 312
QY 335 ATCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 394
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 ATATATGAAGGGCGCTTACAGCTCATGATGATGAGCGCTTCTGTTGTGAAGGA 372
QY 395 CACAGATGATGAGTGTGTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 CACAGATGATGAGTGTGTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 432
QY 455 GACGCGTCTCTATATGCTGACGCTCCCGAGGAACCTTGG 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 GATGATTCCTGTACATGCTGTATGCTCCACAGGAGCGTTCG 476

```

RESULT 6

```

US-10-037-270-707
Sequence 707, Application US/10037270
Publication No. US20030104529A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jiah-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei

```

```

APPLICANT: Tillinghast, John
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: PL-Fl-genes Version 1.0
SEQ ID NO 707
LENGTH: 768
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (109)...(486)
US-10-037-270-707

```

```

Query Match          33.6%; Score 215; DB 14; Length 768;
Best Local Similarity 74.1%; Pred. No. 1.6e-46;
Matches 272; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```

```

QY 132 CCGCGCGAGTCCCTCAGACCGCGCTTCAAGAGCGGGGAGCTTCCGCGCGCGTA 191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 CCGCGACCATGCGCGTGGAGAAAGACCTTCAAGAGCGCGCGCGCGCGCGCGAG 160
QY 192 AGAGGTACAGCATGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 161 AAGATGTCGACATTAATTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 220
QY 252 ACAAGGTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 ACAAGGTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 280
QY 312 TCACATGAGCAGTGTGTCAGATCATCGCGCGCGCGCGCGCGCGCGCGCGCG 371
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 TCACATGAGTGTGTCATCATGATTAATGAGAGGGCGTTACAGCTCATGCTAG 340
QY 372 CCTTCTCTGCGTGTGAGAACCGACAGCAGATGCTGCTGCTATGCTAGCGCT 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 CCTTCTCTGCTGTGAGAACCGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 400
QY 432 TCTAGCAGCAGAGAAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 491
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 TGTATGAGAGTGAAGAAAGATGAAGATGATGCTCTGTACATGCTATGCTCC 460
QY 492 CCTTCGG 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 CGTTCGG 467

```

RESULT 7

```

US-09-925-301-209
Sequence 209, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 209
LENGTH: 2250

```


Best Local Similarity 80.1%; Pred. No. 1.1e-40;
Matches 226; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 217 GCACCCAGCAAAATCCGGTATATCGAGGCTTACAGGAGTGAAGACAGTCCCT 276
DB 537 GCATCCACCAAGATCCAGTATATAGAGCATCAAGGTAAGAGAGTGCCT 478
QY 277 CCGGACAGACCAATTTTGGTCCCGACATGACATGAGCAATGGTCAAGAT 336
DB 477 CCGGACAGACCAATTTTGGTCCCGACATGACATGAGCAATGGTCAAGAT 418
QY 337 CATCCGCGCCGCTGACATGACACCCAGAGGCTTCTGCTGTAACAGCA 396
DB 417 AATTGAAGGCGCCGCTGACATGACACCCAGAGGCTTCTGCTGTAACAGCA 358
QY 397 CAGCATGTAATGTGTCCAGCCCATCGGACATCTACGAGCAGAGAAAGCAGCA 456
DB 357 CAGCATGTAATGTGTCCAGCCCATCGGACATCTTGAAGTACGAGAGAGATGAGAA 298
QY 457 CGGCTTCTCTATGATGCTACGCTCCAGGAAACCTTCG 498
DB 297 CGGCTTCTCTATGATGCTACGCTCCAGGAAACGTTTCG 256

RESULT 10
US-09-864-761-28415
; Sequence 28415, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203

QY 343 GCGCCGCTGACAGCTGACACCCAGCAGGCTTCTTCTGCTGTAACACAGCAT 402
DB 1 GCGCCGCTGACAGCTGACACCCAGCAGGCTTCTTCTGCTGTAACACAGCAT 60
QY 403 GGTGAGTGTGTCCAGCCCATGCGGACATCTACGAGCAGAGAAAGCAGAGCGCTT 462
DB 61 GGTGAGTGTGTCCAGCCCATGCGGACATCTACGAGCAGAGAAAGCAGAGCGCTT 120
QY 463 CCTCTATATGCTTACGCTCCAGGAAACCTTCGCTTCTGA 505
DB 121 CCTCTATATGCTTACGCTCCAGGAAACCTTCGCTTCTGA 163

RESULT 11
US-10-277-156-1
; Sequence 1, Application US/10277156
; Publication No. US20030054399A1
; GENERAL INFORMATION:
; APPLICANT: LADUNGA et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CLO00842 DIV
; CURRENT APPLICATION NUMBER: US/10/277,156
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/708,725
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/243,428
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Human
US-10-277-156-1
Query Match 16.9%; Score 108.4; DB 14; Length 201;
Best Local Similarity 74.7%; Pred. No. 6.7e-19;
Matches 136; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 317 ATGAGCAGTGTGATCATATCCGGCGCTGACGTAACCCAGCAGGCTTC 376
DB 1 ATGAGTACGCTCATCAAGATATTAAGAGGCGCTTACAGCTCATGCTATACAGGCTTC 60
QY 377 TTCTGCTGTGACACGACAGCATGTGATGTGACGAGCCATCGGAGATATAC 436
DB 61 TTCTGCTGTGACACGACAGCATGTGATGTGACGAGCCATCAACCAATCTAGAGTAT 120
QY 437 GAGCAGAGAAAGCAGAGAGCGCTTCCTATATGCTTACGCTCCAGGAAACCTTC 496

Db 121 GAAGAGTGAAGAGATGAGATGATTCCTCTGATCATGCTGTGTGCTTCCAGGAGACGTTT 180
 QY 497 GG 498
 Db 181 GG 182

```

RESULT 12
US-09-864-761-885/C
Sequence: 885, Application US/09864761
Patent No. US20020048765A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 885
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078644.10
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93

```

```

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HL600, SIGNAL = 1.3
US-09-864-761-885

```

Query Match	14.9%	Score 95.6	DB %	Length 474
Best Local Similarity	65.4%	Pred. No. 1.5e-15		
Matches 140	Conservative 0	Mismatches 74	Indels 0	Gaps 0
QY	261	AGAAAGCAGCTGCCCCCTCTGTGCAGACGCTAAATTTTGGTCCCGGAGCAATGCAATATGA	420	
Db	474	AAAGGACGCTTCCTCTCTGSAAMAACCAAGATATTTGTACCTGATACGGTCAATATGA	415	
QY	321	GCGATGGTCAAGATTCATCGCGCGCGCTGTACGTGAAAGCGCGAGGAGGCTTCCTG	480	
Db	414	GTAAAGCTGCTCAAGATTAATTAACAAAGCCCTTTAAAGCTGCATATGTAATCAAGCGCTTCCTG	455	
QY	381	TGCTGGTGAACCAATACAGCATGTGTAGATGTGTCCAGCGCTATTCGGGAGCAATGAATACAT	440	
Db	354	TAGTGTGTAGAGAGACAGCAGCTTGGTAGATTTTCATGCTGATTAATGAAAGTATGAAA	295	
QY	441	ACGAAAGACGAGAGACGGCTTCCTCTATATGCT	474	
Db	294	GTGAAGAGATGAAGAGCAATTCCTGTGATGAT	261	

```

1 RESULT 13
2 US-09-864-761-17665/c
3 Sequence 17665, Application US/09864761
4 Patent No. US20020048763A1
5 GENERAL INFORMATION:
6 APPLICANT: Penn, Sharon G.
7 APPLICANT: Rank, David R.
8 APPLICANT: Hanzel, David K.
9 APPLICANT: Chen, Weisping
10 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
11 FILE REFERENCE: Aecomica-X-1
12 CURRENT APPLICATION NUMBER: US/09/864,761
13 CURRENT FILING DATE: 2001-05-23
14 PRIOR APPLICATION NUMBER: US 60/180,312
15 PRIOR FILING DATE: 2000-02-04
16 PRIOR APPLICATION NUMBER: US 60/207,456
17 PRIOR FILING DATE: 2000-05-26
18 PRIOR APPLICATION NUMBER: US 09/642,366
19 PRIOR FILING DATE: 2000-08-03
20 PRIOR APPLICATION NUMBER: GB 24263,6
21 PRIOR FILING DATE: 2000-10-04
22 PRIOR APPLICATION NUMBER: US 60/236,359
23 PRIOR FILING DATE: 2000-09-27
24 PRIOR APPLICATION NUMBER: PCT/US01/00666
25 PRIOR FILING DATE: 2001-01-30
26 PRIOR APPLICATION NUMBER: PCT/US01/00667
27 PRIOR FILING DATE: 2001-01-30
28 PRIOR APPLICATION NUMBER: PCT/US01/00664
29 PRIOR FILING DATE: 2001-01-30
30 PRIOR APPLICATION NUMBER: PCT/US01/00669
31 PRIOR FILING DATE: 2001-01-30
32 PRIOR APPLICATION NUMBER: PCT/US01/00665
33 PRIOR FILING DATE: 2001-01-30
34 PRIOR APPLICATION NUMBER: PCT/US01/00668
35 PRIOR FILING DATE: 2001-01-30
36 PRIOR APPLICATION NUMBER: PCT/US01/00663
37 PRIOR FILING DATE: 2001-01-30
38 PRIOR APPLICATION NUMBER: PCT/US01/00662
39 PRIOR FILING DATE: 2001-01-30
40 PRIOR APPLICATION NUMBER: PCT/US01/00661
41 PRIOR FILING DATE: 2001-01-30
42 PRIOR APPLICATION NUMBER: PCT/US01/00670
43 PRIOR FILING DATE: 2001-01-30
44 PRIOR APPLICATION NUMBER: US 60/244,687
45 PRIOR FILING DATE: 2000-09-21
46 PRIOR APPLICATION NUMBER: US 09/508,408

```

PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 17665
LENGTH: 155
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078644.10
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
OTHER INFORMATION: NT HIT: U80885.1, EVALUE 2.00e-11
OTHER INFORMATION: SWISSPROT HIT: O41515, EVALUE 2.00e-10
OTHER INFORMATION: EST_HUMAN HIT: AA723823.1, EVALUE 3.00e-72
US-09-864-761-17665

Query Match 11.2%; Score 71.4; DB 9; Length 155;
Best Local Similarity 66.7%; Pred. No. 2.7e-09;
Matches 102; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 267 AGCTGCGGCTCTGCAACAAGCAAGTTTGTGTCGCGGACCATGTCAATGAGCAGT 326
DB 155 AGCTTCTGCTCTGCAAAAACCAAGTAATTTGTACATGATCAACATGAGTAAC 96
QY 327 TGTCAAGATCATCGCGCGCTGACGTGAACCCGAGCGCTTTCTTCTGCTGG 386
DB 95 TCGTCAATATTAGAGGCGCTTAAGCTCGATAGTAATCAAGCCCTTCTTCTGAGTGG 36
QY 387 TGAACGACGACGATGCTGAGTGTGTCACGC 419
DB 35 TGACGAGACGACGCTTGGTGTGATTTTCCATGC 3

RESULT 14
US-09-917-800A-1711
Sequence 1711, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US/09/917,800A
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884

PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 1711
LENGTH: 975
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1_NW_022706
US-09-917-800A-1711

Query Match 11.2%; Score 71.4; DB 10; Length 975;
Best Local Similarity 51.6%; Pred. No. 3e-09;
Matches 215; Conservative 0; Mismatches 196; Indels 6; Gaps 2;

QY 92 TCCCGCGCCGACGAGCCCGGCTGCGCGCCGACCGCGCGCGCATGCTCAGAC 151
DB 53 TCGCTTCCCGCAAGTCTGCGCTCAAGAGCGCGCTCCGCTTCCCGCGCATGAAG 112
QY 152 CGGCTTCAAGCAGCGCGGAGCTTCCCGCGCGCTGTAAGAGAGTACAGATCCGC 211
DB 113 TGGATGTTAAGAGAGACACTCGCTGGAACACAGATCGTGAATCCGCAAGATCAGA 172
QY 212 GACGACAGCCCGAGCAAAATCCCGGTGATCATCGACGCTACAGGAGTGAAGCAGCTG 271
DB 173 GCGAATATCCCGAGCGGCTTCCGCTGATCGTTGAG---AAGTCTCGCTTCAGATT 229
QY 272 CCCGCTCTGACAGACCAAGTTTGTGTCGCGGACCATGTCAATGAGAGTTGGCT 331
DB 230 GTTGACATTGACAMAGAGAGTACTTGTCCATCTGACATCAGTGTGATCATG 289
QY 332 AAGATCATCGCGCGCGCTGAGCTGAACCCGAGCGGCTTCTCTGCTGAGTAC 391
DB 290 TGGATCATCAGAAAGAGATCAGCTTCTTCTGAGAGGCGATCTTCTTGTGTGAGC 349
QY 392 CAGCAGACATGTGTGATGTGTCCAGCCCATCGCGGACATCTACAGCAGAGAAAGAC 451
DB 350 AAGACAGTCCCAAGTCCAGCTTACT--ATGGACAGCTTTACGAGAAAGAAAGAT 406
QY 452 GAGAGCGCTTCTCTATATGCTTACGCTCCAGGAACCTTGGCTTGTGAGCC 508
DB 407 GAAGATGATTCTGTATGTGCGCTTACAGCGAGAGAACCTTTGGCTTGTGAGCC 463

RESULT 15
US-10-205-342-24
Sequence 24, Application US/10205342
Publication No. US20030108906A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alstair
APPLICANT: BrooksBank, Robert
APPLICANT: Pinnock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018198
CURRENT FILING DATE: US/10/205,342
PRIOR APPLICATION NUMBER: US/02-07-24
PRIOR FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 24
LENGTH: 975
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: cDNA: Ganglioside expression factor 2
US-10-205-342-24

Query Match 11.2%; Score 71.4; DB 14; Length 975;

Best Local Similarity 51.6%; Pred. No. 3e-09;
Matches 215; Conservative 0; Mismatches 196; Indels 6; Gaps 2;

```
QY 92 TCCCCGCCCCCAGAGCCCCGGCTGCGGCCCAAGCCGGCCGCGCATGCCCTCAGAC 151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 TCGCTTCGCCGAAGCTCGCGCTCAAAAGCCGGCTCGCTCCGCCGCCATGAA 112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 CGGCGTTTCAGACAGCGCGAGCTTCGCCGACCGCTGTAAAGAGGTACAGCAGATCCGC 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 TGGATGTTTAAAGAGGACCTCGCTGAAACACAGATGCTGGAAATCCGCCAGATCAGA 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 GACCAAGCACCCCAAGCAAAATCCCGGTATCATCGAGCGCTACAAAGGTGAGAACGCGTG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 GCGAAATACCCCGACCGGCTTCGGGTGATCGTTGAG---AAATCTCTGGCTTCAGATT 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 272 CCCGTCTGTGACAGACCAAGTTTGTGTCGCCGACCATGTACATGAGCGAGTTGTC 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 GTTACATTTGACAGAGAGTACTGTGCTCCATCTGACATCACTGTGCTCAGTTCATG 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 AAGATCATCCGCGCGCGCTGAGCTGAACCCCAAGCGGCGCTTCTCTGCTGTGAG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 TGGATCATCAGGAAAAGGATCCAGCTTCTCTGTGAAAGGCCATCTCTTGTGTGGAGC 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 CAGCACAGCATGTGAGTGTGTCCAGGCCCATCGCGACATCTACAGATAGGAGAAAGAC 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 AAGACAGTCCACAGATCCAGCTTACT--ATGGACAGCTTTAGACAGGAAAAGAT 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 452 GAGGACGCTTCTCTATATGTGTACGCTTCCAGAAACCTTGCGCTTCTGAGGC 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 GAAGATGATTTCTGTATGTGCTACAGCGAGAGAACACTTTGCGCTTCTGAGCC 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: August 15, 2003, 10:34:15

Job time : 259 secs

THIS PAGE BLANK (USPTO)